

Constitutional Court Review

Constitutional Court Review



© Juta and Company (Pty) Ltd

1st Floor, Sunclare Building
21 Dreyer Street
Claremont 7708, Cape Town

This journal is copyright under the Berne Convention. In terms of the Copyright Act 98 of 1978, no part of this journal may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording or by any information storage and retrieval system, without permission in writing from the Publisher.

The authors and the publisher believe on the strength of due diligence exercised that this work does not contain any material that is the subject of copyright held by another person. In the alternative, they believe that any protected pre-existing material that may be comprised in it has been used with appropriate authority or has been used in circumstances that make such use permissible under the law.

ISSN: 2073-6215

Design and typesetting by ANdtp Services

Cover design by Drag and Drop

Printed by ????????????????

Contents

Administrative Law

1. Good Reviews, Bad Actors: The Constitutional Court's Procedural Drama 1
Leo Boonzaier
2. The Puzzle of Pronouncing on the Validity of Administrative Action on Review 27
Geo Quinot & P J H Maree
3. The Ebb and Flow of the Application of the Principle of Subsidiarity – Critical Reflections on *Motau* and *My Vote Counts* 43
Melanie Murcott & Werner van der Westhuizen
4. Administrative Action, the Principle of Legality and Deference – The Case of *Minister of Defence and Military Veterans v Motau* 68
Andrew Konstant
5. The Test for 'Exceptional Circumstances' Where an Order of Substitution is Sought: An Analysis of *Trencon* Against the Backdrop of the Separation of Powers 91
Lauren Kohn
6. Clarifying the Exceptional Circumstances Test in *Trencon*: An Opportunity Missed 115
Raisa Cachalia

Affirmative Action

7. Dignity and Equality in *Barnard* 135
Samantha Vice
8. Affirmative Action and Intensity of Review: *South African Police Service v Solidarity obo Barnard* 163
Chris McConnachie

The Zimbabwe Torture Docket Case

9. South Africa's Competing Obligations in Relation to International Crimes 199
Lilian Chenwi & Franziska Sucker
10. Neither Complimentary nor Complementary: *National Commissioner of the South African Police Service v Southern African Litigation Centre and Another* 247
Salim A Nakhjavani

Articles

11. Constitutional Heedlessness and Over-Excitement in the Common Law of Delict's Development 259
Emile Zitzke
12. Against the Interests of Justice: Ignoring Distributive Justice When Certifying Class Actions 291
Khomotso Moshikaro

Case Comment

13. No Place for the Poor: The Governance of Removal in *Zulu* and *SAITF* 321
Irene de Vos & Dennis Webster

Administrative Law

Good Reviews, Bad Actors: The Constitutional Court's Procedural Drama

Leo Boonzaier*

The law must rule, and prevent government officials from acting self-interestedly or arbitrarily. When officials act unlawfully, therefore, their decisions ought to be undone. Indeed, the officials themselves have a duty to set things right in court. But sometimes an official acts self-interestedly or arbitrarily in the very act of undoing her own unlawful decision. What should the court do then? That is the puzzle this article is about.

This puzzle has arisen in two recent Constitutional Court cases.¹ A related puzzle has arisen in a third.² The thrust of the Court's judgments has been to constrain the government's power to have its prior decisions undone by a court. I discuss the puzzle in more detail in Part I, and the Court's solution in Part II. Since the solution has been to erect procedural bars to the official's attempts to undo her own decision, this raises the broader question of the relationship in the Court's jurisprudence between procedure and substance, which is the subject of Part III.

I THE PROBLEM

When and why would we want to tie the government's hands? That is the puzzle I am interested in here. Why, in other words, would we want to say that an official who has gone to court to undo a prior decision that is clearly unlawful should be prevented from doing so? The question may be surprising, since it seems axiomatic that unlawful conduct must be undone. That is the point of

* DPhil candidate, Faculty of Law, University of Oxford. I am grateful to Meghan Finn, who co-hosted the SAIFAC seminar that triggered many of the thoughts in this article, and who continued to provide advice and encouragement thereafter. I also benefited from discussions with Michael Bishop, Cora Hoexter, Chris McConnachie, Khomotso Moshikaro, Melanie Murcott, and other participants at the *Constitutional Court Review* VII conference, as well as with Nurina Ally, Raisa Cachalia, Peter Smith, and Yana van Leeve. Michael Mbikiwa and two anonymous referees gave excellent comments on a subsequent draft. The usual caveats apply.

¹ *Khumalo and Another v MEC for Education, KwaZulu-Natal* [2013] ZACC 49, 2014 (5) SA 579 (CC), 2014 (3) BCLR 333 (CC) ('Khumalo'); *MEC for Health, Eastern Cape and Another v Kirland Investments (Pty) Ltd t/a Eye & Laser Institute* [2014] ZACC 6, 2014 (3) SA 481 (CC), 2014 (5) BCLR 547 (CC) ('Kirland').

² *KwaZulu-Natal Joint Liaison Committee v MEC for Education, KwaZulu-Natal and Others* [2013] ZACC 10, 2013 (4) SA 262 (CC), 2013 (6) BCLR 615 (CC) ('KZN Joint Liaison Committee').

the venerated principle of legality.³ In fact, government officials have a duty, as custodians of the public interest, to undo illegalities.⁴ Why would a court want to stand in their way?

One obvious answer is that undoing a decision upon which people have relied creates uncertainty. That is true, but not my interest here – not least because, as we shall see, it cannot explain the cases. My interest here is a fact more rarely appreciated, which is that officials are capable of acting irrationally, arbitrarily or opportunistically even when they are undoing unlawful decisions. There is no magic to stop the reappearance of the same vices that produced the unlawful decision in the first place. And sometimes even a good review must be stopped if the actor bringing it is behaving badly; or so I shall argue.

The tension this creates can be understood as a tension within the rule of law. The cases I discuss in this article bear this out. Both the majority and dissenting judgments strenuously invoke the rule of law, using it to support verdicts that are diametrically opposed. Some judges say the rule of law gives them reason to allow officials to undo unlawful acts; some say it gives them reason to prevent their undoing. After all, the rule of law paradigmatically includes the principle of legality, which says that unlawful conduct must be undone. And yet officials could flout the rule of law by acting improperly or arbitrarily in the very attempt to undo unlawful conduct. As Alistair Price has recently said, conceptualising the rule of law is a task that South African scholars ‘cannot afford to ignore’.⁵ These tensions within the rule of law, and the Court’s deployment of it, confirm that further academic treatment is needed. No full conceptualisation will be forthcoming in this article, though the polyvalence of the rule of law is an important theme. What I shall discuss is the dilemma for judges that this generates, and the doctrinal tools they have used to manage it.

The dilemma arises against the background of a debate about the status of unlawful administrative acts,⁶ which has important terminological (if not other) relevance and therefore requires brief discussion. The fundamental starting point is that all unlawful administrative acts are void. But this has always been a theoretical rather than a practical proposition. Without the authoritative judgment of a court, one cannot know that an extant administrative decision is, in truth, unlawful and invalid. And it makes little practical difference whether, as a matter of theory, the court’s judgment declares the act void or makes it so. If judicial review is foreclosed for procedural reasons, even an invalid act is ‘as effective ...

³ See, eg, *AllPay Consolidated Investment Holdings (Pty) Ltd and Others v Chief Executive Officer, South African Social Security Agency and Others (No 2)* [2014] ZACC 12, 2014 (4) SA 179 (CC), 2014 (6) BCLR 641 (CC) (*Allpay*) at para 30 (‘Logic, general legal principle, the Constitution and the binding authority of this court all point to a default position that requires the consequences of invalidity to be corrected or reversed where they can no longer be prevented. It is an approach that accords with the rule of law and principle of legality.’)

⁴ See text to note 12 below.

⁵ A Price ‘The Evolution of the Rule of Law’ (2013) 130 *South African Law Journal* 649, 661.

⁶ See generally, L Baxter *Administrative Law* (1984) 355–358; C Forsyth ‘The Theory of the Second Actor Revisited’ (2006) *Acta Juridica* 209, 209–210; C Hoexter *Administrative Law in South Africa* (2nd Edition, 2012) 546–550.

as the most impeccable of orders'.⁷ Moreover, it will often be unjust for a court to undo the effects of an administrative act upon which people have relied. Thus, even as it declares the act invalid, the court has discretion to decide whether it will be left intact.⁸ The upshot is that, on one hand, unlawful administrative acts are axiomatically invalid. On the other, their factual existence must sometimes be sustained. This means that, on top of the notions of lawfulness and validity, a third-level criterion had to be introduced into the ontology of administrative acts, namely whether the act has been 'set aside' by a court. And this third level is, in many ways, the most important. A court may authoritatively declare an administrative action unlawful and invalid, but if it does not set aside the act it is as robustly existent as any other.

In current South African law, it is possible for the official who made the decision to approach a court to set it aside. It was not always so. In the nineteenth century our courts recognised the doctrine of *functus officio*, which prevents officials from reversing their decisions willy-nilly.⁹ Initially, our courts said it followed from this that an official may not even approach a court to have his or her decision set aside;¹⁰ some interested private party had to bring the application. But in the constitutional era, our courts, impelled by the principle of legality, decided otherwise.¹¹ The Supreme Court of Appeal held, in a series of cases starting with *Pepcor*, that officials may apply to court to have their own unlawful decisions set aside.¹² Indeed, officials were held to have a duty to do so, since the state has a special responsibility, as custodian of the public interest, to undo illegalities.

This was an appealing conclusion to reach. But it has given rise to the problems that motivate this article. The basic difficulty is one already mentioned: officials sometimes behave badly even as they try to undo illegalities. The courts need a way to stop those machinations, or at least to avoid becoming an instrument of them. And that need becomes more pressing in times of venality.

Let us start with an extreme case. Imagine some hypothetical country where the President, widely suspected of trying to subjugate the prosecuting authorities, plucks a new head of prosecutions from obscurity. Now suppose that the head turns out to be far more effective and independent than the President anticipated. Fortunately for the President, he finds out that the head had an undisclosed brush with the law

⁷ *Smith v East Elloe Rural District Council* [1956] AC 736, 769–780 (Lord Radcliffe). See also *Harnaker v Minister of the Interior* 1965 (1) SA 372 (C) at 381A–F; *Oudekraal Estates (Pty) Ltd v City of Cape Town and Others* [2004] ZASCA 48, 2004 (6) SA 222 (SCA) ('*Oudekraal*') at para 26.

⁸ *Millennium Waste Management (Pty) Ltd v Chairperson, Tender Board: Limpopo Province and Others* [2007] ZASCA 165, 2008 (2) SA 481 (SCA) at para 23; *Seale v Van Rooyen NO and Others* [2008] ZASCA 28, 2008 (4) SA 43 (SCA) ('*Seale*') at para 13; *Bengwenyama Minerals (Pty) Ltd and Others v Genorah Resources (Pty) Ltd and Others* [2010] ZACC 26, 2011 (4) SA 113 (CC), 2011 (3) BCLR 229 (CC) ('*Bengwenyama*') at paras 81–85.

⁹ See generally, DM Pretorius 'The Origins of the *Functus Officio* Doctrine, with Specific Reference to its Application in Administrative Law' (2005) 122 *South African Law Journal* 832.

¹⁰ *Ibid* 844–845 and 856, citing *Osterloh v Civil Commissioner of Caledon* (1856) 2 Searle 240; *Mining Commissioner of Johannesburg v Getz* 1915 TPD 323.

¹¹ Baxter (note 6 above) at 379–380 (Criticised the earlier decisions for 'lean[ing] too far on the side of the individual and away from the principle of legality').

¹² *Pepcor Retirement Fund and Another v Financial Services Board and Another* [2003] ZASCA 56, 2003 (6) SA 38 (SCA) ('*Pepcor*'); *Municipal Manager: Qaukeni Local Municipality and Another v FV General Trading CC* [2009] ZASCA 66, 2010 (1) SA 356 (SCA); *Ntshangase v MEC for Finance, KwaZulu-Natal and Another* [2009] ZASCA 123, 2010 (3) SA 201 (SCA).

thirty years ago. As it happens, this hypothetical country, like South Africa, treats it as a reviewable irregularity if the President appoints a prosecution head who is not ‘fit and proper’,¹³ and allows government officials, including the President himself, to bring the review application.¹⁴ What would a court do if presented with this application? Would it be bound to do the President’s bidding and undo the clearly illegal decision? And would it make a difference if the President had chosen to appoint this new prosecution head precisely because he knew about the undisclosed misdemeanours, which would hang over him like a Damoclean sword?

Indeed, we do not need such outlandish hypotheticals to make the problem vivid. It is a strategy actually used by government actors in countless recent cases. These are not the heady days of *Pepcor*, when the government was optimistically enlisted to assert the power of the Constitution through judicial review. The government’s litigation strategies have become flagrantly opportunistic and self-interested, using judicial review to evade, rather than to assert, its obligations. It has sought to nullify those obligations by pointing to its own illegal behaviour in incurring them. Once the government allows itself the use of this strategy, it is hard to stop. Under political pressure because you chose a bad building contractor? No problem; just undo the contract by claiming you failed to vet the builder’s tax certificate.¹⁵ Contracted for services you no longer want? Not to worry; have the contract nullified by claiming you concluded it in contravention of your own procurement rules.¹⁶ Made a settlement agreement whose terms now prove inconvenient? Ditto.¹⁷ The possibilities are endless if such arguments are allowed to succeed.

To be sure, they almost never do. The courts have managed to squash them every time. The point is merely that there is an epidemic of these cases which the courts need to salve. Sometimes the cases are bound to fail on the merits, but we want better safeguards than that. Even if the state can contrive a good legal point – even if the President had happened to appoint someone genuinely unfit and improper – arbitrary and self-seeking governance cannot be given free rein. The Court has, fortunately, perceived this fact and has sought to apply the brakes.

This is merely one instance of a more general trend. In the face of growing governmental dysfunction, the Court has newly asserted itself in a range of areas to gird and to goad government and even, where possible, simply to replace

¹³ Section 9(1)(b) of the National Prosecuting Authority Act 32 of 1998, applied in *Democratic Alliance v President of the Republic of South Africa and Others* [2012] ZACC 24, 2013 (1) SA 248 (CC).

¹⁴ One could also imagine that the country allows the President to institute a commission of inquiry on the basis of the possible irregularity, and to suspend the unwanted NDPP while this takes place.

¹⁵ *Country Cloud Trading CC v MEC, Department of Infrastructure Development* [2013] ZASCA 161, 2014 (2) SA 214 (SCA). An appeal on a different point was heard, but dismissed, by the Constitutional Court in *Country Cloud Trading CC v MEC, Department of Infrastructure Development, Gauteng* [2014] ZACC 28, 2015 (1) SA 1 (CC), 2014 (12) BCLR 1397 (CC).

¹⁶ *Gauteng MEC for Health v 3P Consulting (Pty) Ltd* [2010] ZASCA 156, 2012 (2) SA 542 (SCA) (‘3P Consulting’); *Kwa Sani Municipality v Underberg/Himeville Community Watch Association and Another* [2015] ZASCA 24 (‘Kwa Sani’); *Tasima (Pty) Ltd v Department of Transport* [2015] ZASCA 200, [2016] 1 All SA 465 (SCA) (‘Tasima’). The state’s go-to provision here is s 217 of the Constitution, which requires government contracts to be concluded ‘in accordance with a system which is fair, equitable, transparent, competitive and cost-effective’.

¹⁷ *Provincial Government: North West and Another v Tsoga Developers CC and Others* [2015] ZANWHC 36. The Constitutional Court heard an appeal, but refused leave, in *Provincial Government North West and Another v Tsoga Developers CC and Others* [2016] ZACC 9, 2016 (5) BCLR 687 (CC).

it. Most plainly, the Court has developed more robust remedies to hold errant officials to account. Its recent judgment in *Pheko* exhibits a clear determination to put individual officials on the hook in contempt proceedings; they will no longer be allowed to deflect meaningful judicial enforcement by cowering behind the nebulous legal personhood of the state, or by endlessly passing the buck.¹⁸ The Supreme Court of Appeal recently took the dramatic step of imposing a custodial sentence (albeit suspended) on officials who pigheadedly ignored the requirements of administrative law.¹⁹ These moves to craft sharper-edged enforcement powers have been coupled with a greater willingness to make supervisory orders.²⁰ Also notable is the Court's unprecedented willingness to substitute its own decision for that of a government agency in the recent case of *Trencon*.²¹ The message is clear: the government is no longer to be trusted.

II THE SOLUTION

My concern here is how this has played out when the government has tried to have its own decisions reviewed. The synopsis is that the delay bar has come to the fore as a means to scrutinise, and where necessary to thwart, the government's conduct in bringing the review. At common law a judicial review application must be brought within a 'reasonable' time.²² The Promotion of Administrative Justice Act 3 of 2000 (PAJA) rigidifies that requirement significantly by providing that the application 'must be instituted ... not later than 180 days' after the person became aware, or should have become aware, of the decision and the reasons for it.²³ Two rationales are conventionally given for these delay bars: to ensure the finality of administrative decisions, and to curtail prejudice to the other party.²⁴ These certainly have application where the review is sought by a private party, but the delay bars are used for an additional purpose where the government seeks to review its own decision.

¹⁸ *Pheko and Others v Ekurhuleni City (No 2)* [2015] ZACC 10, 2015 (5) SA 600 (CC) ('*Pheko*'). The Court joined the Executive Mayor, Municipal Manager and departmental head of the Ekurhuleni Municipality, as well as the MEC for Human Settlements in the province, so that the Court's earlier supervisory order – and any concomitant contempt proceedings – would be more effective.

¹⁹ *Tasima* (note 16 above) at para 3 of the order.

²⁰ See, eg, *AllPay* (note 3 above); *Pheko* (note 18 above); *Economic Freedom Fighters v Speaker, National Assembly and Others* [2016] ZACC 11, 2016 (3) SA 580 (CC) ('*EFF v Speaker*'). These contrast strikingly with *Minister of Health and Others v Treatment Action Campaign and Others (No 2)* [2002] ZACC 15, 2002 (5) SA 721 (CC) at para 129, where the Court famously declined to issue a structural interdict on the basis that '[t]he government has always respected and executed orders of this Court. There is no reason to believe that it will not do so in the present case.' It is hard to imagine the Court's saying that now.

²¹ *Trencon Construction (Pty) Ltd v Industrial Development Corporation of South Africa Ltd and Another* [2015] ZACC 22, 2015 (5) SA 245 (CC), discussed elsewhere in this volume by Raisa Cachalia and Lauren Kohn respectively.

²² *Wolgroeiens Afslaers (Edms) Bpk v Munisipaliteit van Kaapstad* 1978 (1) SA 13 (A).

²³ Section 7(1). I ignore the complexities introduced by s 7(1)(a). Section 9(1)(b), it should be noted, allows a court to condone a delay of more than 180 days on application by the person concerned. For recent discussion see *Aurecon South Africa (Pty) Ltd v Cape Town City* [2015] ZASCA 209, 2016 (2) SA 199 (SCA) ('*Aurecon*') at paras 13–17.

²⁴ See, eg, *Associated Institutions Pension Fund and Others v Van Zyl and Others* [2004] ZASCA 78, 2005 (2) SA 302 (SCA) ('*Van Zyl*') at para 46. But see note 35 below.

Khumalo, decided in December 2013, is the first Constitutional Court judgment to deal with that scenario.²⁵ Mr Khumalo was appointed to a position in the KwaZulu-Natal Department of Education in April 2004. Other disappointed applicants lodged a grievance shortly thereafter, complaining that he did not meet the requirements stated in the advertisement for the post. The MEC set up a task team, which reported in January 2007 that Mr Khumalo had not met the requirements and should not have been considered. In October 2008, the MEC applied to the Labour Court to have the appointment set aside. Both that court²⁶ and the Labour Appeal Court²⁷ held that, although the MEC had inexplicably delayed in bringing the application, the principle of legality had to be vindicated.²⁸ They allowed the MEC's challenge and set aside the appointment.

But the Constitutional Court unanimously upheld Mr Khumalo's appeal, holding that the MEC should have been non-suited because of her unexplained delay. The majority judgment of Skweyiya J brings out very clearly the fundamental dilemma. The judgment emphasises, from its first paragraph, the importance of the rule of law.²⁹ Yet this foundational principle points in different directions. On one hand, the Court held, the rule of law not only confers standing on the MEC to correct her own department's unlawful decisions, but imposes a duty to do so.³⁰ On the other, 'expeditious and diligent compliance with constitutional duties' is itself 'a requirement of legality'.³¹ Unlawful administrative conduct must be remedied – but remedied according to law:

Because of these fundamental commitments [to the rule of law], a court should be slow to allow procedural obstacles to prevent it from looking into a challenge to the lawfulness of an exercise of public power. But that does not mean that the Constitution has dispensed with the basic procedural requirement that review proceedings are to be brought without undue delay.³²

In this instance there seemed to be a very strong *prima facie* case that Mr Khumalo's appointment was unlawful: even on his own version, 'he ha[d] not gone far enough to show that he met the post's requirements', with the effect that, in terms of the controlling statute, he was 'not entitled to be considered'.³³ Yet the Court nevertheless held that the delay bar trumped the imperative to undo this decision.

The majority judgment considered three reasons for its conclusion. The first was that Mr Khumalo would suffer prejudice if his appointment were to be undone now, nine years after it was made.³⁴ But the Court discountenanced this

²⁵ *Khumalo* (note 1 above). The Court notes the novelty of the situation before it at para 32.

²⁶ MEC, *Department of Education, KwaZulu-Natal v Khumalo and Another* [2010] ZALC 79, 2011 (1) BCLR 94 (LC).

²⁷ *Khumalo and Another v MEC for Education: KwaZulu-Natal* [2012] ZALAC 26, (2013) 34 ILJ 296 (LAC).

²⁸ I gloss over the fate of the several other parties in the case, and consider only that of Mr Khumalo, one of two first-instance respondents who went all the way to the Constitutional Court.

²⁹ *Khumalo* (note 1 above) at para 1.

³⁰ Ibid at paras 35–36.

³¹ Ibid at para 46.

³² Ibid at para 45.

³³ Ibid at paras 60–61.

³⁴ Ibid at para 52.

reason on the basis of the distinction between invalidity and setting aside.³⁵ Any potential prejudice to Mr Khumalo would be accommodated in determining, on the merits, whether his appointment ought to be set aside. It could not, therefore, justify the Court's refusal to hear the merits of the MEC's application at all. The second reason the Court considered was that the MEC's delay meant the lawfulness of the appointment could not be properly determined. Memories inevitably decline, and documentary evidence disappears, over time;³⁶ and in this case the evidence was particularly patchy, the task team's report having noted an 'almost total collapse' of memory even by early 2007.³⁷ The Court did not discountenance this reason, but should have: if the MEC could not discharge the burden of proving the decision was unlawful, then her application would fail on the merits. The paucity of evidence is rarely, therefore, a sufficient reason to non-suit an applicant altogether.

So the third reason is, it seems to me, the real driver of the Court's decision. The reason is that '[t]he MEC ha[d] not sought in any way to explain her delay'.³⁸ Crucially, the lack of an explanation was not conclusive in itself. What mattered was the inevitable inference to which it led:

The fact that the MEC has elected not to account for the delay, despite having had the opportunity to do so at multiple stages in the litigation, can only lead one to infer that she either had no reason at all or that she was not able to be honest as to her real reasons.³⁹

Suspicious about the MEC's motives seem, then, to have been decisive. Was she really trying to vindicate the rule of law, or was she using the illegality opportunistically to remove an opponent? If the former, why not act sooner? Did the supposed collapse of memory within the department seem like a suspicious cover-up? Was it suggestive that, shortly after Mr Khumalo's appointment and before moves were made to oust him, the African National Congress had wrested control of the KwaZulu-Natal government from the Inkatha Freedom Party? Maybe. Or maybe this is all too conspiratorial. But the Court could not know, because the MEC did not tell it. In the absence of an explanation, the uncomfortable inference arose that the MEC was acting either arbitrarily or dishonestly.

What is more, the MEC's motives mattered only because she was a state functionary. 'Had the matter been brought by a private litigant', the Court held, 'this aspect of the test might weigh less heavily'.⁴⁰ Whereas private litigants may invoke their legal rights for whatever purpose they wish, the rule of law requires

³⁵ Ibid at paras 53, 56. This approach differs from that taken in *Van Zyl* (note 24 above) at para 53 and *Gqwetha v Transkei Development Corporation Ltd and Others* [2005] ZASCA 51, 2006 (2) SA 603 (SCA) at paras 21–23, where the potential prejudice of undoing the decision is treated as a reason to refuse to adjudicate the review at all. Importantly, the Constitutional Court held in *Bengwenyama* (note 8 above) that a court must declare any unlawful decision invalid – which requires it to adjudicate the review – and ameliorate any prejudice in deciding whether or not to set the decision aside. This seems effectively to eliminate the prejudice-avoiding function of the delay bar (except, as we shall soon see, where the prejudice relates to the conduct of the litigation itself).

³⁶ *Khumalo* (note 1 above) at para 66.

³⁷ Ibid at para 7.

³⁸ Ibid at para 39.

³⁹ Ibid at para 51.

⁴⁰ Ibid.

government officials to act without arbitrariness or opportunism at all times – even when trying to undo their own unlawful decisions.

Khumalo embodies the dilemma faced by a court that has to decide whether to allow an official to undo her own unlawful decisions. On one hand, undoing an illegality is required by the rule of law. On the other, in undoing an illegality an official may flout the rule of law. This, I have suggested, drew the Court in *Khumalo* into examining the functionary's motives. Was the MEC nobly cleaning up her department, or was she opportunistically using the illegality to get rid of an employee who had fallen out of political favour? Fortunately, the rule against delay – or, more accurately, the test to determine whether the delay ought to be overlooked – gave a ready doctrinal home for these sorts of considerations. Things are not always so easy.

In *Kirland*, the Eastern Cape Department of Health had granted an approval to Kirland in 2007 to build a private hospital.⁴¹ The following year the Superintendent-General (SG) 'withdrew' the approval. Kirland went to court seeking to set aside the withdrawal on the basis that the SG was *functus officio* and bound by the earlier approval. But then the department revealed what had gone on behind the scenes. In 2006, an internal decision had been made by the SG to refuse Kirland's application on the basis that the region was oversupplied. This was never acted upon or communicated to Kirland. Instead, when the SG was away on sick leave, his acting replacement yielded to political pressure (applied, apparently, by the Eastern Cape MEC and the ANC's then provincial chairperson, Stone Sizani) and granted the approval to Kirland. As the Acting SG explained in her affidavit, the department was seen to be 'withholding licences from BEE companies to establish private hospitals', and so refusing Kirland would 'put her in a bad light in the political arena'.⁴²

The MEC argued that the initial 2006 refusal was the extant decision, and so the *functus officio* rule in fact meant the Acting SG had had no power to grant the 2007 approval. But, if not, the 2007 approval was unlawful because it had been given as result of unauthorised dictation, as the Acting SG's affidavit admitted.

Of course, Kirland had not asked for the 2007 approval to be set aside. It did not know about the dubious circumstances underlying the approval until the MEC revealed them in her opposing affidavit; and, in any event, Kirland would have had no interest in undoing the decision that was in its favour – its litigation was aimed precisely at reviving that decision. Crucially, the department, despite arguing that the approval was unlawful, did not counter-apply for it to be declared invalid or set aside – though Kirland had made the point, in its replying affidavit, that in the absence of a counter-application the government's protestations about the approval were futile. Nevertheless, in addition to affirming all of Kirland's arguments and setting aside the withdrawal, the High Court also set aside the approval, agreeing with the department that the Acting SG had unlawfully

⁴¹ *Kirland* (note 1 above).

⁴² Acting SG's affidavit, quoted in *MEC for Health, Province of Eastern Cape and Another v Kirland Investments (Pty) Ltd t/a Eye & Laser Institute* [2013] ZASCA 58, 2014 (3) SA 219 (SCA) ('*Kirland* (SCA)') at paras 8–9.

buckled under political pressure.⁴³ But the court gave no reasons why it was open to it to set aside the approval in the absence of any application for its judicial review. It simply stated, incorrectly, that Kirland had sought its review.⁴⁴

This was the crucial issue on appeal. The Supreme Court of Appeal held (per Plasket AJA) that the High Court had correctly set aside the withdrawal, but had clearly 'erred' in saying Kirland had also sought the setting aside of the approval.⁴⁵ In truth, the court had had 'no jurisdiction to set aside the approvals granted by [the Acting SG] in the absence of either an application or a counter-application in which that relief was sought'.⁴⁶ And, as long as the approval was not set aside, it 'existed in fact and had legal consequences' and had to be treated as valid.⁴⁷ The Supreme Court of Appeal therefore dismissed the government's appeal but upheld a cross-appeal by Kirland against the setting aside of the approval.

A further appeal by the government was decided by the Constitutional Court in March 2014.⁴⁸ The Court split 7:3. The majority (per Cameron J, with Froneman J adding a separate concurrence) agreed with the Supreme Court of Appeal in all relevant respects. The minority (per Jafta J, with Zondo J adding a separate concurrence) held that the propriety of the approval had been adequately ventilated in the High Court papers and that it should be set aside. To do otherwise would allow a clearly unlawful decision to remain in existence on the 'technical basis' that there was no application to have it reviewed,⁴⁹ thus 'plac[ing] form way above substance'.⁵⁰

In response, the majority judgment of Cameron J gave three '[i]nter-related reasons' why the validity of the approval could not fairly be decided in the absence of a formal counter-application.⁵¹ First, only if there were a counter-application – in which Kirland would be the respondent, rather than, as here, the applicant – would Kirland have the procedural right to have its version of the facts preferred,⁵² and be on notice that it needed to dispute seriously the version given in the Acting SG's affidavit, and if necessary refer the matter to trial, to get the order it sought.⁵³ Secondly, Kirland said it had spent R15 000 000 to acquire property in reliance on the approval.⁵⁴ This and other potential prejudice 'should,

⁴³ *Kirland Investment (Pty) Ltd t/a Eye & Laser Institute v MEC for Health, Province of Eastern Cape NO and Others* [2011] ZAECHGHC 78 (Makaula J).

⁴⁴ Ibid at para 8.

⁴⁵ *Kirland* (SCA) (note 42 above) at para 26.

⁴⁶ Ibid at para 27.

⁴⁷ Ibid at paras 22 and 32.

⁴⁸ *Kirland* (note 1 above).

⁴⁹ Ibid at para 43.

⁵⁰ Ibid at para 50.

⁵¹ Ibid at para 70. I organise these issues in a way different from Cameron J.

⁵² This is a consequence of the rule in *Plascon-Evans Paints Ltd v Van Riebeeck Paints (Pty) Ltd* [1984] ZASCA 51, 1984 (3) SA 623 (A).

⁵³ *Kirland* (note 1 above) at para 85. Notably, Jafta J's minority judgment (at para 33) turned both these points against Kirland. He held that the Acting SG's version must be regarded as correct – and therefore the approval as clearly unlawful – because the department was the respondent. And he held that Kirland 'must now live with the consequences of its choice' not to raise a serious dispute of fact or refer the matter to the hearing of oral evidence.

⁵⁴ Ibid at paras 75, 86.

at least, be properly explored and considered before a decision on which the entity in good faith relied is set aside.⁵⁵

Thirdly, and most interestingly for my purposes, Cameron J pointed to the fact that the department's officials had not explained the seven-month delay between the return of the SG and their first attempts to undo the decision.⁵⁶ 'The Court and Kirland', he said, 'are entitled to know what happened in that time'.⁵⁷ In part, Cameron J's point is that if the MEC had brought a formal counter-application, it would have been hit by the rule against delay, and that the MEC could not be allowed to bypass a clear legislative requirement using 'procedural tricks'.⁵⁸ Here the Court cited *Khumalo*.⁵⁹

As in *Khumalo*, the point has a further dimension. Cameron J is not merely asserting procedural propriety for its own sake. The delay bar and concomitant explanation help the Court by allowing it to assess the department's motives in seeking the setting aside of the decision. Cameron J notes that the department's conduct during the seven-month delay 'suggests that at least it was resigned to the approval', and perhaps that it thought the approval was valid, which 'may count against it' when a court adjudicates its application.⁶⁰ There were also suggestions that the department was still under the MEC's thumb and could not do anything about the approval.⁶¹ Perhaps 'political shenanigans' underlay not only the approval but also the MEC's attempt to undo it? Or were the supposed shenanigans invented by the department to try to undo a decision that had become inconvenient? Unlikely – but impossible to rule out, at least in these impaired proceedings.

What is beyond doubt is that the department claimed that the approval was invalid only after it had become embroiled in adversarial proceedings – and once its earlier decision to grant the approval must have become very inconvenient. Prior to that, the department exhibited not the faintest interest in asserting the decision's unlawfulness, and indeed seemed to presuppose that it was valid.⁶² In these circumstances the Court may well have doubted that the department's sole aim, in asking for the approval to be set aside, was nobly to vindicate the rule of law.

So it is easy to understand why the majority was not willing to undo the approval. Even when a decision is unlawful there are, I have been suggesting, deep-rooted reasons not to allow the government to undo it. This is both because

⁵⁵ Ibid at para 86. As Jafta J suggests (at paras 59–61), the potential prejudice is assessed in determining whether an invalid decision should be set aside, and is not a reason to withhold a declaration of invalidity or refuse to adjudicate the matter at all. Compare the discussion of *Khumalo* (note 1 above). But Cameron J's point is that, absent a counter-application, there was insufficient evidence and argument to assess prejudice, since Kirland had not been on notice to address the matter. This distinguishes the case from *Khumalo*, where there was a formal judicial review application and the potential prejudice had been fully ventilated in both lower courts.

⁵⁶ *Kirland* (note 1 above) at paras 70–74.

⁵⁷ Ibid at para 70.

⁵⁸ Ibid at 83.

⁵⁹ Ibid. Froneman J's concurring judgment (at paras 112–114) quotes from *Khumalo* at length.

⁶⁰ Ibid at paras 72–74.

⁶¹ Ibid at para 71.

⁶² Ibid at paras 71–76.

undoing the decision may set back the interests of those who have relied on it, and because government actors can, in seeking to have the decision undone, exhibit the very same self-interest, partisanship, arbitrariness and other vices that made the decision reviewable in the first place.

For this reason, the minority judgments in *Kirland* are disappointing. They fail to spot the dilemma that this article takes as its starting point, and sail benightedly past the reasons that this article has identified against undoing even an unlawful decision. Jafta J's central argumentative strategy is to stress the impropriety of the approval: it was, he says, not only 'unacceptable and disgraceful' but also 'opportunistic[...]', 'exploit[ative]', 'illegitimate',⁶³ 'fraudulent'⁶⁴ and 'corrupt'.⁶⁵ But this is unhelpful. No one sought to defend the decision; the central issue was how to weigh up the importance of correcting illegalities against the difficulties that correction inevitably entails. Rather than resolving that dilemma, Jafta J denies that there is one.

Cameron J gives the lie to Jafta J's reasoning, and notes the importance of the official's motive in undoing the decision, when he says:

It does not assist the debate to point out that what happened in this case seems to have been highly unscrupulous and deplorable. This is because, in the next case, the official who seeks to ignore departmental action may not be acting with pure motives. Though the official here seems to have been on the side of the angels, the risk of vindicating the Department's approach lies in other cases where the revoker may not be acting nobly.⁶⁶

Jafta J also says that a decision so flagrantly unlawful is 'antithetical to the founding values of our constitutional order' and 'inconsistent with the rule of law'.⁶⁷ But this, again, treats the rule of law as nothing more than the principle that unlawfulness must not be allowed, and must be undone when it occurs. That is to say, Jafta J makes the error, cautioned against by Allan, of adopting an analysis of the rule of law that is no more sophisticated 'than the jejune idea of legality'.⁶⁸ He ignores the fact, to which these cases attest, that the government can act antithetically to the rule of law even as it purports to assert legality.

What the cases also show is just how stacked the deck can be in favour of the government. It has all the information about what went on behind the scenes, while the citizen often cannot know that an illegality has occurred. This means that, as in *Kirland*, the government can triumphantly play its joker whenever – and only if – it is convenient. Worse, the pressure is strong to contrive false information that casts doubt on the decision. If the official does this, the private party, as an outsider to the state's inner workings, is poorly placed to refute the allegations. Sometimes, in fact, government officials can serve these ends not by disclosing or inventing information, but by withholding it. Whereas in ordinary judicial

⁶³ Ibid at para 41.

⁶⁴ Ibid at para 44.

⁶⁵ Ibid at para 46.

⁶⁶ Ibid at para 104.

⁶⁷ Ibid at para 47, quoting from *S v Shaik and Others* [2008] ZACC 7, 2008 (5) SA 354 (CC), 2008 (8) BCLR 834 (CC) at para 72 and *South African Association of Personal Injury Lawyers v Heath and Others* [2000] ZACC 22, 2001 (1) SA 883 (CC), 2001 (1) BCLR 77 (CC) at para 4 respectively.

⁶⁸ TRS Allan 'The Rule of Law as the Rule of Reason: Consent and Constitutionalism' (1999) 115 *Law Quarterly Review* 221.

review cases it is the government's aim to show that the decision had the weight of information and deliberation behind it, its precise purpose in these cases is to show the decision to be unlawful. If, as in *Khumalo*, the full background to the decision would be needed to show that the decision was justified, the government can simply withhold it.

True, these points should not be overstated. The information is not entirely within the state's control; it is legally bound to provide the record of the decision when a judicial review application is brought. But there are certainly cases in which there is no realistic prospect that anyone but the decision-makers themselves will have evidence of the alleged unlawfulness. *Kirland* is one such case; *Tasima*, which I discuss below, is another.⁶⁹ Public procurement cases such as these raise particular worries.⁷⁰ The applicable legal regulations are especially detailed and technical; one can find missteps in almost any tender process, and construct a plausible-sounding review ground almost whenever one has an interest in doing so. Where the government is known to be corrupt, suspicions that it has such an interest are ever-present: the procurement process is its primary means of distributing largesse, and having previous procurement decisions undone is an effective way to free up some more.

Indeed, all this might bear on how the government conducts itself when the decision is made in the first place. Why bother to comply with procurement law if that is only going to make your contract harder to undo if it proves inconvenient? And why not wilfully appoint an NDPP with skeletons in his closet so that you can undo his appointment if he shows too much independence? The point should, again, not be overstated; there remain obvious reasons, both legal and prudential, for officials not to flout the law. But their incentives are plainly messier than one would like. Certainly it is true that where the government has the power to create, to reveal, and to undo irregularities, there are difficulties that do not arise where judicial review is brought by a private party, and which must be specially addressed.

Fortunately, *Khumalo* and *Kirland* have already made important moves to do so. First, the majority judgment in *Kirland* rejected the government's invitation to the Court to revisit *Oudekraal*.⁷¹ Instead, the Court resoundingly endorsed that judgment, making clear that the principle at its core – that the government may not simply ignore its own decisions – extends well beyond second-actor cases and admits of no exception, even in cases involving clear illegalities.⁷² Although the Court was deeply divided in *Kirland* itself, the majority judgment's pivotal

⁶⁹ *Tasima* (note 16 above).

⁷⁰ See also the cases cited in notes 15 and 16 above.

⁷¹ Note 7 above.

⁷² This was despite arguments to the contrary in DM Pretorius 'The Status and Force of Defective Administrative Decisions Pending Judicial Pronouncement' (2009) 126 *South African Law Journal* 537, which cites *Seale* (note 8 above) and was approved in *Gardner and Others v Central University of Technology, Free State* [2010] ZALC 75 at para 57.

passage has since been endorsed by a unanimous Court that included *Kirland's* three dissentients.⁷³

Secondly, both *Khumalo* and *Kirland* assert the importance of the delay bars, both at common law and under PAJA. PAJA's exacting 180-day deadline for the institution of proceedings has been criticised for limiting private citizens' access to court.⁷⁴ But, in the case of officials trying to undo their own decisions, its utility is clear. Here the delay bar's stringency is its strength.⁷⁵ It allowed the Court in *Khumalo* to apply the brakes to the MEC's attempt to undo Mr Khumalo's appointment.⁷⁶ Moreover, the enquiry into the MEC's delay afforded an opportunity to scrutinise her motive. In *Kirland*, the delay bar served a similar function, though at one remove. It was not directly applicable, because the MEC had not brought a proper review application; but it was for that very reason that the majority held that she must bring one, so that she would confront the delay bar rather than evade it.⁷⁷ If our fictional President were to apply to court to undo the appointment of his prosecution head, the delay bar would be the only obvious way to stop him.

Could other solutions be devised? One would be to reverse *Pepcor*,⁷⁸ and hold that public authorities simply have no standing to challenge their own decisions. That would be drastic, but, in light of the many problems that *Pepcor* creates, not without merit. Nevertheless, the Court expressly approved *Pepcor* in *Khumalo*, holding that the MEC has standing under the principle of legality to challenge her own decision.⁷⁹ So this drastic solution seems to have been foreclosed.

On the other hand, *Kirland* intriguingly alludes to, and leaves open, the possibility that public authorities lack standing under PAJA.⁸⁰ Yet if the Court believes that to be a viable interpretation, it is wrong. Nothing in the legislation supports that interpretation: to the contrary, PAJA never says or suggests that the applicant for judicial review must have had its own rights or interests affected.⁸¹ Officials are therefore free to bring PAJA applications in the public interest,

⁷³ *EFF v Speaker* (note 20 above) at para 74. On the other hand, Jafta J's lone concurrence a month later in *Nkata v FirstRand Bank Ltd* [2016] ZACC 12, 2016 (4) SA 257 (CC), 2016 (6) BCLR 794 (CC) at paras 186–187 exploits the wiggle room left open by para 103 fn 78 of *Kirland* to hold that a default judgment granted incorrectly 'was a nullity and had no force in law' and can be ignored by a court even without a declaration of invalidity. The breadth of these passages is irreconcilable with *Kirland* (which is nowhere mentioned) and suggests a continued willingness to curtail its precedential effect where possible. The separate judgment of Nugent AJ, in which Cameron J concurs, criticises Jafta J's approach at paras 158–161.

⁷⁴ C Plasket *The Fundamental Right to Just Administrative Action: Judicial Review of Administrative Action in the Democratic South Africa* (PhD thesis, Rhodes University, 2002) 276–285.

⁷⁵ It is true that government can ask for the delay to be condoned: see note 23 above. But the court retains a wide discretion to grant or refuse condonation, considering much the same factors as under the common-law delay bar.

⁷⁶ The majority (per Skweyiya J) used the common-law delay bar, since in its view the claim was one based on the principle of legality, not PAJA. Zondo J used PAJA's delay bar, having disagreed on this point.

⁷⁷ *Kirland* (note 1 above) at para 83.

⁷⁸ Note 12 above.

⁷⁹ *Khumalo* (note 1 above) at para 32.

⁸⁰ *Kirland* (note 1 above) at para 82 fn 43.

⁸¹ For example, s 6(1) says unqualifiedly that '[a]ny person' may institute judicial review proceedings. It contrasts tellingly with the right in s 5(1) to ask for reasons, which is limited to any person 'whose rights have been materially and adversely affected'.

consonantly with their duties recognised in both *Pepcor* and *Khumalo*. Indeed, to hold otherwise would fall foul of the Constitution's generous standing provisions, with which PAJA must be consistently interpreted.⁸² Finally, granting standing to public authorities under the principle of legality, but denying it under PAJA, would aggravate the notorious bifurcation of South African administrative law.⁸³ *Khumalo* must, in short, be taken to have established that administrators have standing under both the common law and PAJA.

A more narrowly tailored solution would be a purpose-built doctrine, analogous to the 'clean hands' doctrine in private law, that requires the state to prove its bona fides to the court before it is entitled to the relief it seeks. This would be an express doctrinal articulation of the imperatives I have discussed in this article, and would allow courts to consider transparently the factors I have claimed they are considering anyway. And – unlike PAJA's delay bar, of course – it would allow the courts to control mischievous reviews brought by government officials even within 180 days. It is, in short, the ideal solution.

The courts may nevertheless prove reluctant to adopt it. This is, in part, for the very reason that it requires them to acknowledge openly, but on incomplete evidence, that they doubt the government's bona fides – so much so that they are willing for that reason to deny relief. In any event, the common law is famously slow to bring forth grand new doctrines. Sometimes courts really must 'forge new tools' to achieve what is needed.⁸⁴ More commonly, however, they use (or misuse) whatever existing tools suffice to solve the instant case. And in this context, as we have seen, the courts usually have to hand a tool of that kind, namely the delay bar. We await a major case where the government responds with anywhere near the promptness required to make the delay bar inapplicable. And if it does so, it may well be a case where no check on its power is appropriate; for it is usually the very fact of the delay, and the poor explanation for it, that makes an inference of bad faith possible. So in most cases – or until someone brings a constitutional challenge to the delay bar⁸⁵ – it is likely to prove an adequate, if makeshift, tool. Or so it proved, at any rate, in *Khumalo* and *Kirland*.

⁸² Section 38. See also s 33 (in fulfilment of which PAJA was enacted) read with s 39(2).

⁸³ See, eg, C Hoexter 'The Principle of Legality in South African Administrative Law' (2004) 4 *Macquarie Law Journal* 165; I Currie 'What Difference Does the Promotion of Administrative Justice Act Make to Administrative Law?' (2006) *Acta Juridica* 330; L Kohn 'Our Curious Administrative Law Love Triangle: The Complex Interplay between the PAJA, the Constitution and the Common Law' (2013) 28 *Southern African Public Law* 22.

⁸⁴ This phrase, now regularly used by the Court, originates from *Fose v Minister of Safety and Security* [1997] ZACC 6, 1997 (3) SA 786 (CC), 1997 (7) BCLR 851 (CC) at para 69.

⁸⁵ Plasket (note 74 above) argued in 2002 that the 180-day rule was probably unconstitutional and raised the possibility of a court challenge. In *Moblomi v Minister of Defence* [1996] ZACC 20, 1997 (1) SA 124 (CC), 1996 (12) BCLR 1559 (CC) and *Brümmer v Minister for Social Development and Others* [2009] ZACC 21, 2009 (6) SA 323 (CC), 2009 (11) BCLR 1075 (CC) the Constitutional Court declared similar time-bar provisions unconstitutional on the basis that they violated the constitutional right of access to court. But those provisions were significantly more restrictive than PAJA's, since neither afforded an opportunity for a court to condone a delay, and the one in *Brümmer* gave a window of just 30 days – far more restrictive than PAJA, whose limit of 180 days the Court used as a 'yardstick' (para 67) and indeed saw fit to emulate in its reading in (at para 89). A period of 180 days is also used in s 3(2) of the Institution of Legal Proceedings Against Certain Organs of State Act 40 of 2002 and is considerably more generous than England and Wales's 90-day cut-off for the institution of a judicial review: see CPR 54.5(1). These considerations may help to explain why no one has yet been willing

And so it proved in *Tasima*, an important recent judgment of the Supreme Court of Appeal.⁸⁶ The Department of Transport's highly suspect attempt to evade the consequences of its flagrant violation of due process, and of the contract it had concluded with the applicant, was thwarted in large part by Brand JA's appropriately strict application of PAJA's delay bar. Having quoted from *Kirland* at length, he expressed grave doubts about why the department had waited five years to seek the review of its decision.⁸⁷ The department, in its own, less pleasing echo of *Kirland*, insisted its delay should be overlooked because of the serious 'fraud and corruption' tainting the decisions it sought now to undo.⁸⁸ But Brand JA was sceptical. First, why was the department raising the allegations of corruption only now?⁸⁹ Worse, the allegations seemed contrived: the department's conduct of the litigation showed it 'ha[d] little, if any, confidence in its ability to establish these allegations if they were to be properly challenged in a court of law'.⁹⁰ So it was, once again, the delay bar – and the opportunity to interrogate the state's motives that it affords – that allowed the court to deflect an illegitimate review attempt.

But closing the door to belated judicial review applications by public authorities has caused them to push at another. In several recent cases, including *Tasima*, public authorities have tried to skirt the strict procedural requirements attaching to judicial review applications by bringing 'collateral challenges', which were recognised in South African law in *Oudekraal*.⁹¹ The Supreme Court of Appeal has moved decisively to shut this door, ruling in *Merafong* that collateral challenges are not available to organs of state.⁹² The municipality's reliance on a collateral challenge there appeared to be a nakedly opportunistic attempt to avoid PAJA's delay bar, and was rightly rejected. In *Kirland* itself, in fact, one of the MEC's main arguments had been that, supposing she had failed to meet the procedural requirements of a judicial review application, she was entitled to bring a collateral challenge – an argument that the Court did not entertain.⁹³ And, of course, it was the main message of *Kirland* (and indeed of *Khumalo*) that when the government

to bring a constitutional challenge to PAJA's delay bar even though it has been a decisive obstacle to numerous litigants.

Even if a challenge were successful, it would matter what the 180-day rule is replaced with. If a longer period were specified, courts would have little power to stop judicial review applications brought within it. But if the more flexible common-law position were restored, in terms of which judicial review proceedings must be brought within a 'reasonable' time, courts would be given greater scope to clamp down on the state in appropriate cases.

⁸⁶ *Tasima* (note 16 above). In both *3P Consulting* and *Kwa Sani* (note 16 above) the SCA noted the alternative argument that the government's challenge to its own decision was time-barred, but did not have to decide this; the government's inept application failed on the merits: *3P Consulting* (note 16 above) at paras 13–14; *Kwa Sani* (note 16 above) at paras 31–33.

⁸⁷ *Tasima* (note 16 above) at paras 27 and 31.

⁸⁸ *Ibid* at para 32.

⁸⁹ *Ibid* at para 37.

⁹⁰ *Ibid* at para 36.

⁹¹ *Oudekraal* (note 7 above), applying *Boddington v British Transport Police* [1999] 2 AC 143.

⁹² *Merafong City v AngloGold Ashanti Ltd* [2015] ZASCA 85, 2016 (2) SA 176 (SCA) ('*Merafong*') at para 17. Strictly speaking, the judgment says only that an organ of state cannot raise a collateral challenge *against another organ of state*. But it was confirmed, and applied in a dispute between an organ of state and a private party, in *Tasima* (note 16 above) at paras 26–27. See also *Kwa Sani* (note 16 above) at para 14.

⁹³ See the applicant's heads of argument at paras 43–45. The Court's judgment does not mention this even though, if it thought a collateral challenge were permissible, that would have had clear decisional relevance.

tries to plot an easier procedural route it must, ipso facto, be stopped. The twice-stated view of the Supreme Court of Appeal is that the prohibition on the state's bringing collateral challenges is entailed by *Kirland*.⁹⁴ One expects the Constitutional Court to keep this door shut too.⁹⁵

III PROCEDURE AND SUBSTANCE

The theme of the previous sections has been that government officials must sometimes be held to their prior decisions even when they are unlawful, and that procedural exactitude from judges is needed to achieve this. Cameron and Froneman JJ were our protagonists, insisting on scrupulous adherence to legal process. By contrast, Jafta and Zondo JJ took what I argued was an unreasonably lax approach. But what, then, are we to say about the case of *KZN Joint Liaison Committee*,⁹⁶ decided shortly before *Khumalo*, and the subject of an excellent note by Cora Hoexter?⁹⁷

In 2008 the KwaZulu-Natal MEC for Education (of *Khumalo* fame) issued a notice to independent schools in the province setting out the value of their state subsidy (their 'approximate funding levels') for the 2009–10 financial year. In May 2009, when the first tranche had already fallen due, the department sent a letter advising the schools to expect a 30% cut in their subsidies. When the first two tranches were paid, in mid-July, they were indeed 30% short. The applicant, an association of independent schools in the province, entered into negotiations with the department, but in the end the subsidies for the year remained, on average, 30% less than those in the 2008 notice. The association applied to the High Court for an order enforcing the 'promises' in that notice.⁹⁸ The High Court dismissed the application on the basis that, even assuming the notice was intended to create contractual obligations, an obligation to pay an 'approximate' amount of money is too vague to be enforced. Both the High Court and the Supreme Court of Appeal denied leave to appeal.

The Constitutional Court heard the appeal, however, and the conduct of the association's case generated controversy. Importantly, the association had not brought a judicial review application. This, it became clear, was a major obstacle to its now relying on administrative-law grounds. The association had not complied with PAJA's formalities, so no record of the department's decision was before the Court; and, even if it had complied, the association would (at least absent a good

⁹⁴ *Kwa Sani* (note 16 above) at paras 14–17; *Merajong* (note 92 above) at paras 15–17. Admittedly the breadth of these precedents is questionable. *Pace* the Supreme Court of Appeal, the fact that the applicant is an organ of state should not necessarily disqualify it from a collateral challenge: if one organ of state tries to coerce another into compliance with a clearly reviewable decision, it may, at least in some circumstances, be reasonable to allow the latter to await enforcement, safe in the knowledge that it will be able to raise a collateral challenge. Nevertheless, the result in *Merajong* was surely correct.

⁹⁵ Editor's note: The Constitutional Court has, since the article was finalised, decided the appeals in both *Merajong* and *Tasima*. See *Merajong City Local Municipality v AngloGold Ashanti Ltd* [2016] ZACC 35, 2017 (2) 211 (CC), 2017 (2) BCLR 182 (CC); *Department of Transport and Others v Tasima (Pty) Ltd* [2016] ZACC 39, 2017 (2) 622 (CC), 2017 (1) BCLR 1 (CC).

⁹⁶ Note 2 above.

⁹⁷ C Hoexter 'The Enforcement of an Official Promise: Form, Substance and the Constitutional Court' (2015) 132 *South African Law Journal* 207.

⁹⁸ *KwaZulu-Natal Joint Liaison Committee v Member of the Executive Council, Department of Education, KwaZulu-Natal (KZP)* unreported case no 9594/2010 (26 September 2011).

explanation) have been non-suited by the delay bar.⁹⁹ To fix these mistakes would require the association to reinstitute in the High Court, which would drag out the litigation considerably. Counsel for the association said it had no time or money to do this and thus emphatically disavowed reliance on administrative law, saying the case must be decided, 'do or die', by the question whether the 2008 notice created an enforceable promise.¹⁰⁰

In the first place, then, the association's case was that the department's undertaking amounted in law to a contract. But, in the view of almost all members of the Court, there were insurmountable problems with that argument. The case therefore became about a rather different question, namely whether the association was entitled to succeed on some other basis – even though, as the association had acknowledged, it had failed woefully to comply with the procedural requirements of its obvious alternative cause of action, judicial review. As in *Kirland*, therefore, the basic legal difficulty was the same: in the absence of a formal review application, and in the face of PAJA's delay bar, could the Court get to the merits? In the end, the Court split 6:4. As in *Kirland*, Cameron and Froneman JJ respectively wrote the two judgments for the majority, and Jafta and Zondo JJ were in the minority.

But then comes the twist. Quite unlike in *Kirland*, Cameron and Froneman JJ were on the side saying procedural strictures could be bypassed, and the merits of the decision adjudicated; and Jafta and Zondo JJ found themselves insisting that the Court may adjudicate only the case that had been brought with full procedural propriety. This raises an intriguing puzzle, as Hoexter points out.¹⁰¹ What explains the about-face between *Kirland* and *KZN*? Why do Cameron and Froneman JJ insist on procedural niceties in *Kirland*, but happily overlook them in *KZN*? After all, the applicant in *KZN*, on its own version, was deliberately trying to evade PAJA's delay bar and the legislation's other procedural requirements – trying, in other words, to do exactly what Cameron and Froneman JJ prohibited, and Jafta and Zondo JJ countenanced, in *Kirland*. So why do the pairs now swap roles?

The start of the explanation is, of course, that in *Kirland* the party seeking judicial review was a public authority, and in *KZN* it was a private party. Hoexter elaborates:

Most members of the court [including Cameron and Froneman JJ] would hold public authorities to a higher procedural standard than other litigants – an attitude that does not seem inappropriate in view of the state's greater resources and the special duties on it, such as constitutional duties to 'respect, protect, promote and fulfil' rights (s 7(2)) and to perform all constitutional obligations 'diligently and without delay' (s 237).¹⁰²

There is undoubtedly something to this explanation, which derives, after all, from Cameron J's own statement, in a crucial passage in *Kirland*, that '[g]overnment is not an indigent or bewildered litigant, adrift on a sea of litigious uncertainty, to whom the courts must extend a procedure-circumventing lifeline'.¹⁰³

⁹⁹ *KZN Joint Liaison Committee* (note 2 above) at para 31.

¹⁰⁰ *Ibid* at para 34.

¹⁰¹ Hoexter (note 97 above) at 216.

¹⁰² *Ibid* at 217.

¹⁰³ *Kirland* (note 1 above) at para 82.

But there is surely more to it. The fundamental point is not about the competence the respective judges expect of different legal teams. The fundamental point derives from what I have been suggesting throughout this article about state power, and the courts' responsibility to constrain it. The point is simply this: in *Kirland*, insisting on procedural rectitude resulted in the department's being held to its prior decision; in *KZN*, by contrast, doing so would have allowed the department to escape.

As we have seen, in both *Kirland* and *KZN* the key issue was whether the government should be held, against its will, to its own prior decision. The government's position in both was that it should be allowed to wriggle out. The difference is that in *Kirland* procedural strictures were helpful in restraining the government's attempt to undo its decision, but in *KZN* they threatened to do the opposite. In *Kirland* the MEC sought the judicial review of her department's extant decision; success in the review would have allowed the MEC to evade that decision; and therefore insisting that the review be conducted with full procedural propriety ensured that she was bound by her prior decision. In *KZN*, the schools applied to hold the MEC to her undertaking; here, success in the application would have *prevented* the MEC from evading that decision; and therefore the way to hold her to it was to overlook the application's procedural shortcomings.

Seen in this light, *KZN* and *Kirland* are easily reconciled. The two pairs of judges have not reversed roles at all. Their judgments seem inconstant if one focusses upon their attitude to procedural rules, but exactly what one would expect if one focusses upon the object to which those procedural rules were directed: holding (or releasing) the government from its prior decisions. On that issue, both pairs of judges have been unyielding. Cameron J's majority judgment in *Kirland* raised serious doubts about the MEC's attempt to review her decision, and made clear that these must significantly counterweigh the imperative to undo illegalities; and that same concern about the government's motives in withdrawing its undertaking would explain his stance in *KZN*. Jafta and Zondo JJ had no similar worries in *Kirland*; they thought the rule of law was the 'jejune' principle of legality, and nothing more. That same sanguinity about the government's revising its prior decisions would explain why they were unmoved to see the association's application scuppered in *KZN*.

So it is not the case that the only, or primary, contrast here is between the principle of legality and scrupulous adherence to procedure for the sake of certainty and finality in litigation.¹⁰⁴ The richer contrast is between the principle of legality and the fact that the rule of law requires the government always to act

¹⁰⁴ See Hoexter (note 97 above) at 218: ('The rule of law is a multi-faceted concept, and one must surely accept that any relaxation of procedural requirements is bound to erode it in one way or another – irrespective of the immediate beneficiary of the relaxation, and even if there is no discernible prejudice to the other party. The real challenge ... is to decide when and in what circumstances the erosion of one aspect of the rule of law, such as legal certainty, may be justifiable in order to uphold some other aspect of the rule of law, such as legality. But there will always be some sacrifice.') Also see Froneman J's own statement in *Kirland* (note 1 above) at para 114 ('The law does not allow us to uphold the rule of law while at the same time circumvent and undermine it [by circumventing PAJA's delay bar]. In the long run, shortcuts of this kind will erode the rule of law as one of the foundational values of our Constitution.' In light of the position he took in *KZN Joint Liaison Committee*, which plainly circumvents and undermines PAJA's delay bar, one would hope a deeper-going explanation is available.)

for legitimate reasons, both when it makes decisions and when it tries to undo them. Procedural rules are a way of ensuring the latter – sometimes. If the real end is holding the government to its prior decisions, we should not be surprised that the procedural rules are modulated according to their usefulness as a means.¹⁰⁵ Another way of putting the point is this: Yes, indubitably the government's special status matters when procedural rules are applied to it. But that is not because of its special status *qua* litigant; it's not about its extra 'resources' or whether its legal team is 'bewildered'. Its special status is that the rule of law requires it to be held to its prior decisions in ways that private parties are not. If exacting procedural rules are helping to achieve this, expect judges to deploy them strictly. If they aren't, don't. That is what *Kirland* and *KZN* show. Cameron and Froneman JJ's treatment of the procedural rules varies dramatically between the two cases, but only because of the underlying continuity between them, namely the need to keep the government honest.¹⁰⁶

As we have seen, Jafta and Zondo JJ's treatment of the procedural rules also varied dramatically between the two cases, but in the opposite direction. Where Cameron and Froneman JJ saw pressing reasons not to allow even an unlawful decision to be set aside, Jafta and Zondo JJ saw none. But we want to know why. An answer is of course latent in what I already said. The crucial reason against allowing the government to undo its unlawful decisions is its questionable motive in doing so. The compulsion that judges feel to apply the brakes depends, then, on how wary they are of state abuse of power. So, the thought continues, Jafta and Zondo JJ do not apply the brakes as firmly because they are slower to doubt the government's motives in undoing its prior decisions. Whereas Cameron and Froneman JJ view the power of the state with suspicion, and disrupt it in favour of the citizenry, Jafta and Zondo JJ are enablers, willing to empower the state and to prevent litigants from throwing spanners in its works.

How far does this attitude extend? Is it a general truth that Jafta and Zondo JJ are unconcerned about abuse of power by the state? That would certainly resonate with the popular view of Jafta and Zondo JJ (and arguably of their frequent ally

¹⁰⁵ This means-end point should not be overstated. In *Khumalo* (note 1 above) and *Kirland* (note 1 above), at least, the Court squarely addressed what procedures the government should be made to follow before it could be allowed to undo its own decision. That was the main issue. Deciding it required no legal gymnastics, and could be accommodated in the delay enquiry. *KZN Joint Liaison Committee* (note 2 above), on the other hand, was a singular, and badly litigated, case in which the clear doctrinal tool for controlling the government's liability, substantive legitimate expectations, was foreclosed for procedural reasons: see Hoexter (note 97 above) at 224–229 and M Murcott 'A Future for the Doctrine of Substantive Legitimate Expectations? The Implications of *KwaZulu-Natal Joint Liaison Committee v MEC for Education, KwaZulu-Natal*' (2015) 18 *Potchefstroom Electronic Law Journal* 3133. A more creative – and controversial – solution was therefore called for, which clearly did take procedural liberties for the sake of the merits.

¹⁰⁶ There are, to be sure, significant differences between the approaches of Cameron J and of Froneman J in *KZN Joint Liaison Committee*: see Hoexter (note 97 above) at 210–211. Here and in what follows, my analysis contrasts one pair of judges, Cameron and Froneman JJ, with another pair, Jafta and Zondo JJ. Quite obviously, this risks suppressing the disagreements between the members of each of those pairs, and a more complete analysis would bring those out. My incipient suggestions in that direction are relegated to footnotes.

Mogoeng CJ) as the Court's conservative, pro-government members.¹⁰⁷ But is it true? The number of cases in which Jafta J has found in favour of the government for procedural reasons that a majority of the Court found unpersuasive is certainly striking.¹⁰⁸ In addition to *Kirland* and *KZN*, one can point to *Oriani-Ambrosini v Sisulu*¹⁰⁹ and *Mazibuko v Sisulu*,¹¹⁰ two politically charged cases in which Jafta J's procedural objections meant he would have left Parliament's rules, which the majority found unconstitutional because they disadvantaged minority parties, untouched. And in the last two 'school cases', *Welkom* and *Rivonia*, Jafta and Zondo JJ held that officials from the provincial Department of Education had the power to intervene in the decisions of school governing bodies in circumstances where the majority of the Court held it impermissible.¹¹¹ In *Rivonia*, the decisive reason for Jafta J's dissent was procedural: the majority's finding against the department was, he believed, not open to the Court because it had not been properly pleaded.¹¹² Finally, one might cite *SAPS v Barnard*, in which Jafta J adopted a very restrictive attitude to procedure, the effect of which was to defeat Ms Barnard's attempt to challenge the SAPS's affirmative action decision.¹¹³

These well-known judgments have earned Jafta and Zondo JJ their reputation as the Court's 'stickler[s] for process' and the stalwarts of its 'procedurally conservative camp'.¹¹⁴ Certainly they adopt formalist¹¹⁵ techniques – including, notably, the minute parsing of the first-instance pleadings – which their colleagues would find embarrassing.¹¹⁶ I have not yet mentioned *Maphango and Others v Aengus Lifestyle Properties (Pty) Ltd*, another judgment of Cameron J (supplemented by a concurrence

¹⁰⁷ See, eg, R Calland *The Zuma Years* (2013) ch 11; E McKaiser 'Shame on those two ConCourt judges' *The Star* (7 October 2013); N Tolsi 'Applause for Mogoeng's judicial cadenza' *Mail & Guardian* (18 October 2013). As is well known, Mogoeng's public reputation has shifted: he is now regarded as an important counterweight to the power of the government, especially as a result of his judgment in *EFF v Speaker* (note 20 above).

¹⁰⁸ Some of these are discussed in Hoexter (note 97 above) at 214.

¹⁰⁹ *Oriani-Ambrosini v Sisulu, Speaker of the National Assembly* [2012] ZACC 27, 2012 (6) SA 588 (CC), 2013 (1) BCLR 14 (CC). Zondo J was not yet on the Court.

¹¹⁰ *Mazibuko NO v Sisulu and Others NNO* [2013] ZACC 28, 2013 (6) SA 249 (CC), 2013 (11) BCLR 1297 (CC). Here Zondo J concurred in Jafta J's dissent.

¹¹¹ *Head of Department, Department of Education, Free State Province v Welkom High School and Others* [2013] ZACC 25, 2014 (2) SA 228 (CC), 2013 (9) BCLR 989 (CC); *MEC for Education, Gauteng Province, and Others v Governing Body, Rivonia Primary School and Others* [2013] ZACC 34, 2013 (6) SA 582 (CC), 2013 (12) BCLR 1365 (CC) ('*Rivonia*').

¹¹² *Rivonia* (note 111 above) at para 105.

¹¹³ *South African Police Service v Solidarity obo Barnard* [2014] ZACC 23, 2014 (6) SA 123 (CC), 2014 (10) BCLR 1195 (CC) ('*Barnard*'). The joint judgment written by Cameron and Froneman JJ and Majiedt AJ agreed that Ms Barnard's application had to fail, though not for procedural reasons. It decided the case on its merits, which it thought were finely poised.

¹¹⁴ Hoexter (note 97 above) at 214.

¹¹⁵ This term is hardly ideal, for reasons classically expressed in F Schauer 'Formalism' (1988) 97 *Yale Law Journal* 509. Yet its use in the contemporary South African debate that I am about to enter is too entrenched to avoid. Its meaning in this context is, I think, tolerably clear. It is more precise, at least in one respect, than the usages Schauer lamented: it refers to rigidity in the application of legal rules of a specific kind, namely rules of procedure.

¹¹⁶ See, eg, *KZN Joint Liaison Committee* (note 2 above) at paras 151–154; *Rivonia* (note 111 above) at paras 92–99 and 108; *Khumalo* (note 1 above) at paras 78 and 81–90. Even in *Kirland* (note 1 above), where Jafta and Zondo JJ adopted the more permissive approach to procedure, a fixation upon the wording of the parties' first-instance affidavits was a hallmark of their dissents: see paras 37, 48, 124–127, 139 and 143–147.

by Froneman J) from which Zondo J (with Mogoeng CJ and Jafta J concurring) vigorously dissented on the basis that the remedy ordered by the majority had not been pleaded¹¹⁷ – the Mogoeng Court's first significant procedure-based split, and one that has already spawned a burgeoning literature.¹¹⁸ These formalist techniques, as we have seen, often manifest in ways that benefit state litigants.

On the other hand, there is a series of judgments – generally less well-known – in which Jafta J has displayed a pliability in the face of procedural rules that is unmatched by even his avowedly 'anti-formalist' colleagues.¹¹⁹ There is *Kirland*, of course, as well as his lofty remarks elsewhere that courts should show 'flexibility' in the application of procedural rules where their 'mechanical application' would 'lead to an injustice',¹²⁰ and not waste time on 'mere formalities which are not dispositive of a real dispute'.¹²¹ But I am thinking of those cases where his procedural flexibility leads him to find, more surprisingly, *against* the state. Witness, for example, his remarkable willingness to chart a lone course, circumventing what seemed unanswerable procedural irregularities to reach the merits, and strike down a state policy, in *Sali v National Commissioner of the SAPS*.¹²² And note his radical interventionism in setting aside the report of a state commission that a majority of his colleagues thought unimpeachable in *Bapedi Marota Mamone v Commission on Traditional Leadership Disputes and Claims*, doing so on the basis of a novel argument raised by none of the parties – a move Khampepe J's majority judgment thought most unfair.¹²³ In the earlier case of *Tulip Diamonds*, Jafta J took a strikingly permissive approach to the rules of standing, accordingly reached the

¹¹⁷ [2012] ZACC 2, 2012 (3) SA 531 (CC), 2012 (5) BCLR 449 (CC). See in particular paras 103–104, 109–110 and 113.

¹¹⁸ J Fowkes 'Managerial Adjudication, Constitutional Civil Procedure and *Maphango v Aengus Lifestyle Properties*' (2013) 5 *Constitutional Court Review* 309; M Dafel 'On the Flexible Procedure of Housing Eviction Applications' (2013) 5 *Constitutional Court Review* 331.

¹¹⁹ The phrase is again Hoexter's. Hoexter (note 97 above) at 210 and 215. She uses it to describe Cameron and Froneman JJ, and cites the latter's comments in *KZN Joint Liaison Committee* at para 79 and his 'Legal Reasoning and Legal Culture: Our "Vision" of Law' (2005) 16 *Stellenbosch Law Review* 3.

¹²⁰ *PFE International and Others v Industrial Development Corporation of South Africa Ltd* [2012] ZACC 21, 2013 (1) SA 1 (CC), 2013 (1) BCLR 55 (CC) at para 31.

¹²¹ *Dengetenge Holdings (Pty) Ltd v Southern Sphere Mining & Development Co Ltd and Others* [2013] ZACC 48, 2014 (5) SA 138 (CC), 2014 (3) BCLR 265 (CC) at para 136.

¹²² [2014] ZACC 19, 2014 (9) BCLR 997 (CC). Cameron J wrote the majority judgment which was adamant that the merits could not be fairly decided. Unusually, even Zondo J could not agree with Jafta J.

¹²³ [2014] ZACC 36, 2015 (3) BCLR 268 (CC) ('*Mamone*'). Here, too, Zondo J disagreed with Jafta J. Both judges have been willing elsewhere to decide cases on the basis of issues entirely unheralded by any of the parties or the lower courts. For example, in *Cool Ideas 1186 CC v Hubbard and Another* [2014] ZACC 16, 2014 (4) SA 474 (CC) Jafta J's dissent, concurred in by Zondo J, would have decided the matter on the basis that the parties' contract was invalid because it contravened legislation, even though the parties had formally agreed that the litigation was to be decided on the contrary assumption. And in *Horn and Others v LA Health Medical Scheme and Another* [2015] ZACC 13, 2015 (7) BCLR 780 (CC) Zondo J would have decided the matter on the basis that the respondent's alleged obligation to pay pension benefits to the appellants, its former employees, had been transferred to their new employer by automatic operation of s 197 of the Labour Relations Act 66 of 1995, even though the respondent itself had disavowed that possibility as a 'non-issue' in its founding papers and both parties repeated that stance before the Constitutional Court. Here, at least, unlike in *Cool Ideas*, the Court asked after the hearing for written submissions on this novel point. But that would have offered little comfort to the third party, not involved in the litigation at all, to whom Zondo J's judgment said these liabilities had been transferred. Fortunately, a majority of the Court held that to impose obligations on a third party *mero motu* and without its joinder 'should not be done' (para 35).

merits of the applicant's judicial review – where a majority of his colleagues, and both of the lower courts, would not – and decided it nonchalantly against the government.¹²⁴

So describing Jafta and Zondo JJ as the Court's procedural sticklers tells only half the story. Something more complex is going on. We need to explain why they veer from sticklers to libertines. As I have suggested, a full explanation is not going to be the judges' attitudes to procedure itself – for why, then, would their attitudes seem so inconstant? – but in the causes that procedure can be made to serve; and the operative cause, to which the Court's judges have divergent loyalties, is constraining the powers of the state. But there, too, we have already encountered several wrinkles, in the form of cases in which Jafta and Zondo JJ relaxed procedural rules to the state's detriment. A further wrinkle, and an intriguing counterpoint to *Kirland*, is *S v Nabolisa*, where the procedural rectitude of the majority judgment of Jafta J, now back in stickler mode and insisting on a full, formal application to cross-appeal by the state, gave individuals important protections from the power of the prosecution authorities.¹²⁵ Indeed, even in *Kirland* itself, Jafta J's willingness to castigate the former MEC shows the need for a suitably fine-grained appraisal.¹²⁶ One must also account, finally, for Jafta J's concurrences in a range of highly politically charged judgments of the Court that found unflinchingly against the national government.¹²⁷

To be sure, one would not expect a single criterion to determine a judge's entire judicial philosophy. And thus far I have spoken, simplistically, as though the state

In any event, Jafta J and Zondo J's approach in *Mamone*, *Cool Ideas* and *Horn* sits most uncomfortably with their fierce dissents in *KZN Joint Liaison Committee* (note 2 above), where, although the majority undoubtedly decided the case on a basis of the Court's, not the applicant's, making, the novelty was largely nominal, not substantive; and where the applicant's only reason for disavowing the basis on which the Court ultimately decided the matter was that it did not want to have to reinstitute in the High Court – a wish the Court respected. Their approach also sits uncomfortably with *Kirland* (note 1 above) at para 147, where Zondo J holds that the Court can and must ignore applicable legislative provisions if they have not been raised by the parties.

¹²⁴ *Tulip Diamonds FZE v Minister for Justice and Constitutional Development and Others* [2013] ZACC 19, 2013 (10) BCLR 1180 (CC). This time he was joined by his regular allies, Nkabinde and Zondo JJ.

¹²⁵ *S v Nabolisa* [2013] ZACC 17, 2013 (8) BCLR 964 (CC). Or perhaps the real kicker is that, in the earlier case of *S v Bogaards* [2012] ZACC 23, 2012 (12) BCLR 1261 (CC), Jafta J's dissent would (partly on the basis of misgivings about the way the case was pleaded) have *denied* an almost identical protection to an applicant who had been convicted of harbouring the so-called Boeremag terrorists. In both cases, Zondo J agreed with Jafta J.

¹²⁶ See text to note 65 above.

¹²⁷ *Democratic Alliance v President of the Republic of South Africa and Others* [2012] ZACC 24, 2013 (1) SA 248 (CC), 2012 (12) BCLR 1297 (CC); *Minister of Police and Others v Premier of the Western Cape and Others* [2013] ZACC 33, 2014 (1) SA 1 (CC); *National Commissioner of the South African Police Service v Southern African Human Rights Litigation Centre and Another* [2014] ZACC 30, 2015 (1) SA 315 (CC), 2014 (12) BCLR 1428 (CC); *Helen Suzman Foundation v President of the Republic of South Africa and Others* [2014] ZACC 32, 2015 (2) SA 1 (CC), 2015 (1) BCLR 1 (CC); *EFF v Speaker* (note 20 above). If we are engaging in this crude totting up, then it is worth mentioning Cameron J's concurrences in pro-executive decisions like *South African Reserve Bank and Another v Shuttleworth and Another* [2015] ZACC 17, 2015 (5) SA 146 (CC) (where his frequent ally, Froneman J, wrote a powerful dissent) and his unwillingness to overlook procedural propriety in *Zulu and Others v eThekweni Municipality and Others* [2014] ZACC 17, 2014 (4) SA 590 (CC), 2014 (8) BCLR 971 (CC) (again, Froneman J could not agree). Of course, even in *KZN Joint Liaison Committee* itself Froneman J's 'anti-formalism' was notably more far-reaching than Cameron J's: see note 105 above.

is a monolith. In truth it has diverse parts, and it would be quite appropriate for the judges to take different attitudes to each. For example, does Jafta J, in accordance with traditional separation of powers doctrine, apply a light touch to the political branches, but readily hold the criminal justice system and the bureaucracy to account? That would go some way to explaining his deference to MECs in *Rivonia* and *Kirland*, and to Parliament in *Oriani-Ambrosini* and *Mazibuko*; and his relative willingness to intervene against the police in *Sali*, the prosecutors in *Nabolisa* and the commissioned lawyers in *Mamone*. Or is it that Jafta J is keen to free the government of legal fetters only where he thinks the goal it is pursuing is particularly important, or particularly liable to derailment by interest groups: affirmative action in *Barnard*, liberalising access to privileged schools in *Rivonia*? I shall not take these questions further here. I am content to say only that in *Kirland* and *KZN Cameron and Froneman JJ* clearly perceived a need to restrain public power in a way that Jafta and Zondo JJ did not.

What also cannot be disputed is the sheer number of recent cases that have been dividing the Court on procedural lines. The latest, and one of the most striking, is *My Vote Counts*.¹²⁸ Here Jafta and Zondo JJ did not sit. Nevertheless, a 7:4 majority ducked a question of considerable public interest – whether Parliament has an obligation to enact legislation requiring political parties to disclose their private funding – on the singularly unpersuasive basis that the applicant NGO, *My Vote Counts*, had opted for the wrong procedural route.¹²⁹ The dissenting judgment of Cameron J (with Froneman J and two others concurring) cut through these proceduralist contrivances and found comprehensively in favour of *My Vote Counts*. He would therefore have issued an order that Parliament was in breach of its constitutional obligations. On that momentous issue the majority, of course, had nothing to say; its sole concern was that the applicant had not jumped through what it described as certain ‘procedural hoops’.¹³⁰

That procedure so regularly obscures the merits, even at super-appellate level, is worrying. Still worse is the suspicion that the merits are the true source of the rift in the Court, which the endless procedural contestations merely mask. And the

¹²⁸ *My Vote Counts NPC v Speaker of the National Assembly and Others* [2015] ZACC 31, 2016 (1) SA 132 (CC) (*‘My Vote Counts’*).

¹²⁹ This is not the place for a full engagement with the disappointing majority judgment; the following synopsis, which largely repeats criticisms made by the minority judgment, must suffice. The nub of the majority’s reasoning was that the applicant, *My Vote Counts*, ought to have brought its case by means of a constitutional challenge to the pertinent legislation that already exists, namely the Promotion of Access to Information Act 2 of 2000 (PAIA). But *My Vote Counts* had made clear that, in its view, PAIA was constitutionally valid and indeed essential to realising the constitutional right of access to information. It was therefore very odd to insist that *My Vote Counts* attack that legislation’s constitutional validity. The principle of subsidiarity, on which the majority heavily relied (paras 160–183), did not in truth assist it: *My Vote Counts*, far from subverting or bypassing PAIA, had explained why PAIA was inapplicable and inapposite to the issue of political party funding; Parliament had duly responded to this charge in detail. For the same reason, what was to be gained by forcing *My Vote Counts* to reinstitute in the High Court? That application would raise precisely the same issues about PAIA’s sufficiency that had already been argued at length before the Constitutional Court – and which, indeed, the High Court had already decided, consonantly with *My Vote Counts’* argument, in *Institute for Democracy in South Africa and Others v African National Congress and Others* [2005] ZAWCHC 30, 2005 (5) SA 39 (C).

¹³⁰ *My Vote Counts* (note 128 above) at para 175.

veneer is wearing very thin. As in *Barnard*, so in *My Vote Counts*: ‘the overwhelming impression created by the majority judgment is a Court seizing on an easy way out of a difficult task’.¹³¹ Or is it that the majority seized on an easy way of *doing* a difficult task? Did it believe that the applications of Ms Barnard and My Vote Counts deserved to be dismissed, and that slamming the procedural door was a decisive way to achieve this? Whatever the correct impression, it is not a pleasing one. Elsewhere, the Court seems to have retreated to rather dubious points of procedure precisely so that it can obtain unanimity rather than a nasty merits-based split: Michael Mbikiwa and I have written that the Court’s unconvincing remittal of an important but highly controversial substantive question in *H v Fetal Assessment Centre*¹³² raises this worry.¹³³

The point I have been labouring is that procedure is being subordinated, in various ways, to the judges’ views of the merits. Is it too discomfiting to say that procedural rules are being used as a mere means to a concealed end? Has the argument of this paper been too realist, too deflating, in treating doctrine as a smokescreen for what is really driving judicial decisions? Maybe for some;¹³⁴ but I think this conclusion would be too quick. Partly, this is because it is unwise to deny that judges are doing something because we think they ought not to be doing it. The ostrich-in-the-sand approach is especially dangerous given what is clearly a profound cleavage in the Court that manifests along procedural lines. This cleavage is, shortly put, *the* story of the Court’s last few years. If the Court’s approach to procedural rules has become a key battleground, we need to understand the opposing factions.

Some might say, voicing a cognate complaint, ‘Well, what did you expect?’ They will say the damaging inconsistencies in a riven Court’s application of procedural rules are the foreseeable consequence of its deliberate retreat from formal reasoning into reasoning that is substantive, discretionary and value-

¹³¹ C McConachie ‘Affirmative Action and Intensity of Review: *South African Police Service v Solidarity obo Barnard* (2015) 7 *Constitutional Court Review* 163. He also quotes C Albertyn ‘Adjudicating Affirmative Action within a Normative Framework of Substantive Equality and the Employment Equity Act – An Opportunity Missed? *South African Police Service v Solidarity obo Barnard*’ (2015) 132 *South African Law Journal* 711, 716 (the majority performs an ‘oddly formalistic side-step’ to avoid the main issues).

¹³² [2014] ZACC 34, 2015 (2) SA 193 (CC), 2015 (2) BCLR 127 (CC).

¹³³ L Boonzaier & M Mbikiwa ‘The Constitutional Court in Harms’ Way: A Response’ (2015) 4 *South African Law Journal* 769, 777–778. Pierre de Vos has made a similar suggestion in response to the Court’s dismissal of *De Lange v Methodist Church and Another* [2015] ZACC 35, 2016 (2) SA 1 (CC): see P de Vos ‘Constitutional Court avoids decision on sexual orientation discrimination by Methodist Church’ *Constitutionally Speaking* (24 November 2015), available at <http://constitutionallyspeaking.co.za/constitutional-court-avoids-decision-on-sexual-orientation-discrimination-by-methodist-church>. This seems to me uncharitable; to overcome that application’s very serious procedural irregularities would have been a bridge too far. Nevertheless, De Vos’s perception that the Court is using procedural reasons in a way that is not ‘honest’ is itself telling.

¹³⁴ South African legal scholarship rarely analyses court behaviour from anything approaching a ‘political science’ perspective. The major exception is T Roux *The Politics of Principle: The First South African Constitutional Court, 1995–2005* (2013). The mixed reception to Roux’s methodology is well illustrated by the two reviews of the book in (2014) 131 *South African Law Journal* 704 and 707: Jeremy Gauntlett is scathing; Stephen Ellmann is reverential.

based. Cautions have been sounded against this trend for decades now.¹³⁵ 'We told the Court not to get onto the slippery slope towards arbitrary, ad-hoc, even disingenuous applications of rules that should be adhered to strictly', they might say. 'When courts try too hard to do justice in a particular case, we only end up worse off.'

The critique should, I think, be taken seriously.¹³⁶ But two responses are apt – in addition to the obvious one that refraining from doing justice in particular cases has its own patent cost. First, it is far from clear that the uncertainties and divisions in the Court's approach are the result of a lack of fidelity to existing procedural rules. To the contrary, we have already noted several cases, such as *Maphango* and *Rivonia*, where it was a minority's obsessive and unaccountable fealty to the words used in the pleadings that made things fraught and divisive. The uncertainty and division in these cases is not a consequence of the flexible approach, in other words, but of a recalcitrant minority's preference for formalism.

True, I have stressed that this tells only half the story. When that minority has yawned to the opposite, procedurally unchained extreme this has equally been the cause of deep rifts. Yet the point remains that those rifts cut across the distinction between procedural conservatism and procedural creativity. Neither is a solvent of bad judgement. Minute parsing of the pleadings can be just as productive of senseless and surprising outcomes as the application of flexible tests. Although the Court as a whole is often criticised for its willingness to engage in substantive reasoning at the expense of lawyerly fidelity to formal rules, most of its judges manage to reach a fairly comfortable consensus on procedural questions even as they do so. The point can be stated bluntly: subtract Jafta and Zondo JJ from the Court, and a coherent, predictable and more or less unanimous approach to procedural questions starts to emerge.¹³⁷

The challenge, then, is to understand why those two judges do what they do. Unfortunately, they offer us little help, rarely articulating the purposes they take to underlie the rules and principles they deploy (or choose not to deploy), or the reason why they parse the parties' pleaded case like a statute in one case and shrug it off happily in another, why they insist on full procedural rectitude in one case and disavow it in the next. To be fair, it would be quite wrong to say that Jafta and Zondo JJ are the only offenders here: though they are most frequently in dissent, many of these charges can be levelled at the Court's other judges, who also adopt apparently inconsistent attitudes to procedure that are poorly substantiated. Their

¹³⁵ See, in increasing order of indignation, A Cockrell 'Rainbow Jurisprudence' (1996) 12 *South African Journal on Human Rights* 1; S Woolman 'The Amazing, Vanishing Bill of Rights' (2007) 124 *South African Law Journal* 762; J Lewis 'The Constitutional Court of South Africa' (2009) 125 *Law Quarterly Review* 440; LTC Harms 'The Puisne Judge, the Chaos Theory and the Common Law' (2014) 131 *South African Law Journal* 3.

¹³⁶ See, in addition to Cockrell and Woolman – (note 134 above), H Scott 'The Death of Doctrine? Private Law Scholarship in South Africa Today' in J Basedow, H Fleischer & R Zimmermann (eds) *Legislators, Judges and Professors* (2016) 223.

¹³⁷ This is a further reason to worry particularly about the recent judgments in *Barnard* (note 113 above) and *My Vote Counts* (note 128 above), where procedure divided the Court even apart from the contribution of Jafta and Zondo JJ.

prior decisions are rarely mentioned, let alone distinguished.¹³⁸ Undoubtedly the solution lies, in part, in more astute debate within and without the Court about its adjudicatory role.¹³⁹ But it also lies, I have suggested, in understanding the judges' convictions about the merits, to which procedure is frequently subordinate. That is a large task, to which I have contributed only modestly, if at all, in this article.

But the outlook is not entirely dreary. Sometimes the Court has spoken with one voice even while striking out in entirely novel procedural directions. The most justly celebrated case is *Informal Traders*, where the unanimous Court heard, and granted, an urgent interim application that had leap-frogged the High Court to restrain draconian conduct by the Johannesburg Metro Police.¹⁴⁰ Another example is *Sarrabwitz v Maritz NO and Another*, a very poorly litigated application that all judges agreed merited an unusually flexible approach (though they differed on the detail).¹⁴¹

Here, then, is a more optimistic view about the Court's approach to procedure, and a second response to those who would despair when we notice some space appearing between a court's behaviour and its doctrine. When a judge departs from a common-law rule, it is not anarchy. It is legal development. Doctrines can be revised to better capture the reasons on which they are based. I have explained why I think there is strong pressure on judges to regulate, and sometimes to hinder, the government's attempts to undo even unlawful decisions. The reason, to repeat, is that the government can behave badly even as it undoes a bad decision. Recognising that this is a new challenge, not comfortably accommodated within existing doctrine, was the lesson of Part I. That it is not easily accommodated is not, of course, a reason for judges to ignore it. The truth that judicial decisions should comport with doctrine need not mean that judges should never depart from the existing law. Often they should. It is only that, when they do, their reasons for doing so must be made clear, and that doctrine must be made to catch up. That this can be done was the lesson of Part II. And that academics' recognising the space between doctrine and outcomes is a help, not a hindrance, to improving both – well that, I hope, was the lesson of the whole article.

¹³⁸ For discussion of the Court's general reticence about its own precedents, see J Brickhill 'Precedent and the Constitutional Court' (2010) 3 *Constitutional Court Review* 79; T Ngcukaitobi 'Precedent, Separation of Powers and the Constitutional Court' (2012) *Acta Juridica* 148.

¹³⁹ By far the best body of work on this that I know is J Fowkes 'How to Open the Doors of the Court – Lessons on Access to Justice from Indian PIL' (2011) 27 *South African Journal on Human Rights* 434; J Fowkes 'Civil Procedure in Public Interest Litigation: Tradition, Collaboration and the Managerial Judge' (2012) 1 *Cambridge Journal of International and Comparative Law* 235; Fowkes (note 117 above).

¹⁴⁰ *South African Informal Traders Forum and Others v City of Johannesburg and Others* [2014] ZACC 8, 2014 (4) SA 371 (CC), 2014 (6) BCLR 726 (CC).

¹⁴¹ [2015] ZACC 14, 2015 (4) SA 491 (CC), 2015 (8) BCLR 925 (CC).

The Puzzle of Pronouncing on the Validity of Administrative Action on Review

Geo Quinot*
P J H Maree†

I INTRODUCTION

Two cases decided by the Constitutional Court in 2014 highlighted one of the puzzles of administrative law under the constitutional order. The various judgments in *MEC for Health, Eastern Cape v Kirland Investments (Pty) Ltd*¹ and *AllPay Consolidated Investment Holdings (Pty) Ltd v Chief Executive Officer of the South African Social Security Agency (No 2)*² raised the question of the relationship between a finding that administrative action is unconstitutional and hence invalid, and the setting aside of that administrative action.

At common law the setting aside of administrative action has always been regarded as a discretionary remedy. This approach has not been seriously interrogated under the democratic constitutional regime. However, the framework within which judicial review of administrative action now takes place under the Constitution of the Republic of South Africa, 1996 differs dramatically from that of the common law. The question thus emerges whether the traditional approach to setting aside administrative action can be reconciled with the new constitutional framework.

In what follows, we grapple with this issue by looking at the remedy of setting aside in administrative law. We thus focus exclusively on an order by the court following a successful application for judicial review of administrative action, to the effect that the impugned administrative action ceases to exist. The central question is how that remedy is to be understood in terms of the Constitution. The exact remedy at issue here carries different labels under the different regimes, which are not always consistently used. At common law it was called setting aside administrative action. Under the constitutional regime it is at times referred to as consequential relief following a declaration of invalidity. We are less concerned about the label attached to the remedy and more with the substance of it. Whatever

* Professor, Department of Public Law, Stellenbosch University.

† Post-doctoral fellow, Department of Public Law, Stellenbosch University.

¹ *MEC for Health, Eastern Cape and Another v Kirland Investments (Pty) Ltd t/a Eye & Laser Institute* [2014] ZACC 6, 2014 (3) SA 481 (CC), 2014 (5) BCLR 547 (CC) ('Kirland').

² *AllPay Consolidated Investment Holdings (Pty) Ltd and Others v Chief Executive Officer, South African Social Security Agency and Others (No 2)* [2014] ZACC 12, 2014 (4) SA 179 (CC), 2014 (6) BCLR 641 (CC) ('Allpay 2').

it is called, it is the order, which in substance has the result that the impugned administrative action ceases to exist, that we are interested in.

We start by setting out how the common law framed judicial review of administrative action and the remedy of setting aside. We then move on to consider what the Constitution states about judicial review and remedies before turning to the Court's treatment of the issue in *Allpay 2* and *Kirland*.

While the question that we grapple with in this contribution certainly has particular practical implications, as our analysis of the cases will aim to show, we are concerned first and foremost with doctrine. In our view it is essential to develop a coherent doctrinal approach to questions such as what an appropriate, effective remedy should look like in order to achieve the aim of administrative justice under the Constitution.

II THE COMMON LAW

The starting point of the inquiry is the common law, where it has long been the position that the remedy of setting aside following review of administrative action is within the court's discretion to grant or to withhold.³ This is regardless of the court's finding on the reviewability of the impugned administrative action. In other words, whether the administrative action is reviewable, ie whether a ground of review is found to exist, and whether the administrative action should be set aside as a consequence of the finding of reviewability, have long been held to be separate questions. This goes back to one of the foundational judgments for judicial review of administrative action in South African common law, namely that of Innes CJ in *Johannesburg Consolidated Investment Co Ltd v Johannesburg Town Council*, where he stated:

Whenever a public body has a duty imposed upon it by statute, and disregards important provisions of the statute, or is guilty of gross irregularity or clear illegality in the performance of the duty, this Court may be asked to review the proceedings complained of and set aside or correct them.⁴

A positive answer to the former question, that is of reviewability, is not decisive of the latter question of setting aside.

The discretionary nature of setting aside follows from the discretionary nature of the entire review procedure. As Corbett J stated in *Harnaker v Minister of the Interior*, 'review under the common law' is 'an inherent jurisdiction exercised by the Court'.⁵ The granting or refusal of an order setting aside an irregular administrative action seems to rest on considerations of both principle and practice. In the context of English law, Wade and Forsyth state that while the rule of law would insist on granting such an order, 'the denial of a remedy sometimes serves the public interest', and they give as examples instances where such a refusal is necessary to avoid administrative chaos or to protect the interests of innocent

³ *Oudekraal Estates (Pty) Ltd v City of Cape Town and Others* [2004] ZASCA 48, 2004 (6) SA 222 (SCA) ('*Oudekraal*') at para 36; L Baxter *Administrative Law* (1984) 678, 713.

⁴ 1903 TS 111 at 115 (emphasis added).

⁵ 1965 (1) SA 372 (C) at 380.

third parties.⁶ The authors of *De Smith's Judicial Review* echo the principled position that granting the order must be the default position, but note that fairness and justice may dictate a refusal in a given case.⁷ In relation to South African common law, Baxter observes that '[i]t is not easy to discern a clear set of principles upon which this discretion will be exercised'.⁸

Whatever the exact justification for this remedial discretion may be, at common law the remedy to set aside (*certiorari* to quash) has been discretionary from at least the mid-seventeenth century,⁹ and it is fair to say that the position was well settled in South African common law prior to the adoption of the democratic Constitution. This position was authoritatively stated by the Supreme Court of Appeal in *Oudekraal* in an oft-quoted passage in which the court declared:

[A] court that is asked to set aside an invalid administrative act in proceedings for judicial review has a discretion whether to grant or to withhold the remedy. It is that discretion that accords to judicial review its essential and pivotal role in administrative law, for it constitutes the indispensable moderating tool for avoiding or minimizing injustice when legality and certainty collide.¹⁰

Even though this judgment was handed down in 2004, a decade into the new constitutional era, it was essentially decided in terms of common law. The court relied exclusively on English authority for the above-quoted paragraph, particularly the House of Lords decision in *Wandsworth London Borough Council v Winder*¹¹ and *De Smith*.¹²

III THE CONSTITUTIONAL ERA

While *Oudekraal* was decided with primary reliance on common-law authority and can thus be viewed as a statement of the position at common law, the fact that it was decided well into the constitutional era confirms that the common-law position has been generally accepted up to now. This is borne out by a line of judgments from the Supreme Court of Appeal confirming the discretionary nature of the remedy of setting aside in administrative law. For example, in *Chief Executive Officer, SASSA v Cash Paymaster Services (Pty) Ltd* the court declared that '[c]onsiderations of public interest, pragmatism and practicality should inform the exercise of a judicial discretion whether to set aside administrative action or not',¹³ and in *Judicial Service Commission v Cape Bar Council* the same court held that 'even if an administrative decision is challenged and found wanting, courts

⁶ HWR Wade & CF Forsyth *Administrative Law* (11th Edition, 2014) 597.

⁷ The Rt Hon the Lord Woolf, Professor Sir Jeffrey Jowell QC, Professor Andrew Le Sueur et al *De Smith's Judicial Review* (7th Edition, 2013) 979.

⁸ Baxter (note 3 above) at 713, and see further at 713–718 where he discusses 'some of the more dominant considerations which influence the courts in the exercise of their discretion'.

⁹ Lord Justice Bingham 'Should Public Law Remedies be Discretionary?' 1991 *Public Law* 64, 65.

¹⁰ *Oudekraal* (note 3 above) at para 36 (footnotes omitted).

¹¹ [1985] AC 461 (HL), [1984] 3 All ER 976 (HL).

¹² SA de Smith, H Woolf & JL Jowell *Judicial Review of Administrative Action* (5th Edition, 1995).

¹³ *Chief Executive Officer, South African Social Security Agency, and Others v Cash Paymaster Services (Pty) Ltd* [2011] ZASCA 13, 2012 (1) SA 216 (SCA) at para 29 (footnotes omitted).

still have a residual discretion to refuse to set that decision aside'.¹⁴ At times the court has, however, sought to circumscribe this discretion by stating, for example in *Eskom Holdings Ltd v New Reclamation Group (Pty) Ltd*, that an applicant who succeeds in showing the existence of a reviewable irregularity '[o]rdinarily . . . would be entitled to call for the [administrative action] to be set aside',¹⁵ thus suggesting that setting aside may be viewed as the default remedy.

When comparing these judgments with the common-law position as set out in the previous section, it becomes evident that that position has largely endured under the Constitution. While setting aside was the default remedy when administrative action was found irregular, the reviewing court could refuse it within its discretion. In *New Reclamation*¹⁶ the court furthermore read the Promotion of Administrative Justice Act 3 of 2000 (PAJA) to confirm the discretionary approach to setting aside as expressed in *Oudekraal*. Section 8 of the PAJA gives the reviewing court the power to grant 'any order that is just and equitable, *including* orders . . . (c) setting aside the administrative action' (emphasis added). This formulation seems indeed to support the view of setting aside as a discretionary remedy.

At this point the reader may justifiably ask: So what is the issue? The constitutional regime seems to align perfectly with the common law on this point, as recognised in both legislation and case law. There may thus be no puzzle here at all. However, in our view the issue emerges when one compares the common-law position with what is set out in s 172(1) of the Constitution, which reads:

- (1) When deciding a constitutional matter within its power, a court—
 - (a) must declare that any law or conduct that is inconsistent with the Constitution is invalid to the extent of its inconsistency; and
 - (b) may make any order that is just and equitable, including—
 - (i) an order limiting the retrospective effect of the declaration of invalidity; and
 - (ii) an order suspending the declaration of invalidity for any period and on any conditions, to allow the competent authority to correct the defect.

Unlike Brand JA who, without more, stated in *JSC v Cape Bar Council* that 'this common-law principle . . . is confirmed in effect by s 172(1) of our Constitution',¹⁷ it is not immediately obvious to us that s 172(1) in fact allows for the discretionary nature of setting aside.

The starting point of the puzzle is the recognition that all applications for judicial review of administrative action are now founded on s 33 of the Constitution, regardless of whether the section is in fact expressly raised in the application. That

¹⁴ *Judicial Service Commission and Another v Cape Bar Council and Another* [2012] ZASCA 115, 2013 (1) SA 170 (SCA) ('*JSC v Cape Bar Council*') at para 13.

¹⁵ *Eskom Holdings Ltd and Another v New Reclamation Group (Pty) Ltd* [2009] ZASCA 8, 2009 (4) SA 628 (SCA) ('*New Reclamation*') at para 11.

¹⁶ *Ibid* at para 9.

¹⁷ *JSC v Cape Bar Council* (note 14 above) at para 13.

much is borne out by the judgments in *Pharmaceutical Manufacturers Association*,¹⁸ *Bato Star*¹⁹ and *New Clicks*.²⁰

In *Pharmaceutical Manufacturers* the Court held that the common law does not continue to provide a basis for judicial review of administrative action separate and parallel to that set out in the Constitution.²¹ The Constitution is now the sole basis for judicial review of administrative action, and the common law has been 'subsumed' under it.²² Importantly, the Court held that there is only 'one system of law . . . shaped by the Constitution, which is the supreme law, and all law . . . derives its force from the Constitution'.²³

In *Bato Star* the Court added the PAJA to this structure of judicial review by holding that the cause of action for judicial review of administrative action will now ordinarily arise from that statute.²⁴ The authority of the PAJA rests in turn on the Constitution, and on s 33 in particular.²⁵ The Court further explained that given this relationship between the PAJA and the Constitution, 'matters relating to the interpretation and application of PAJA will . . . be constitutional matters'.²⁶

Finally, in *New Clicks* Chaskalson CJ rejected the approach of both the courts below, which had held that there was a basis for judicial review of administrative action under s 33 independent from the PAJA.²⁷ Chaskalson CJ held that the PAJA was intended to and in fact does 'cover the field', with the result that a litigant cannot go behind it and rely on s 33 directly when bringing an application for judicial review of administrative action.²⁸ This view was supported by Ngcobo J, who held that it is impermissible either for a litigant to argue or a court to decide a review application directly with reference to s 33 without taking the PAJA into account.²⁹ The PAJA was intended to give effect to s 33, and that is the way in which judicial review of administrative action in terms of the requirements of administrative justice set out in s 33 must be argued. Reliance on the PAJA is thus reliance on s 33 itself.

A finding of reviewability under the grounds of review in s 6(2) of the PAJA thus amounts to a finding that s 33 has not been complied with. This brings us

¹⁸ *Pharmaceutical Manufacturers Association of South Africa and Another: In re Ex parte President of the Republic of South Africa and Others* [2000] ZACC 1, 2000 (2) SA 674 (CC), 2000 (3) BCLR 241 (CC) ('*Pharmaceutical Manufacturers*').

¹⁹ *Bato Star Fishing (Pty) Ltd v Minister of Environmental Affairs and Tourism and Others* [2004] ZACC 15, 2004 (4) SA 490 (CC), 2004 (7) BCLR 687 (CC) ('*Bato Star*').

²⁰ *Minister of Health and Another NO v New Clicks South Africa (Pty) Ltd and Others (Treatment Action Campaign and Another as Amici Curiae)* [2005] ZACC 14, 2006 (2) SA 311 (CC), 2006 (1) BCLR (1) (CC) ('*New Clicks*').

²¹ *Pharmaceutical Manufacturers* (note 18 above) at para 33.

²² *Ibid.*

²³ *Ibid* at para 44.

²⁴ *Bato Star* (note 19 above) at para 25.

²⁵ *Ibid.*

²⁶ *Ibid.*

²⁷ *New Clicks* (note 20 above) at para 93, referring to the majority judgment of the High Court in *New Clicks South Africa (Pty) Ltd v Tshabalala-Msimang and Another NNO; Pharmaceutical Society of South Africa and Others v Tshabalala-Msimang and Another NNO* 2005 (2) SA 530 (C), [2005] 1 All SA 196 (C) and to the unanimous judgment of the Supreme Court of Appeal in *Pharmaceutical Society of South Africa and Others v Tshabalala-Msimang and Another NNO; New Clicks South Africa (Pty) Ltd v Minister of Health and Another* [2004] ZASCA 122, 2005 (3) SA 238 (SCA).

²⁸ *New Clicks* (note 20 above) at paras 95–96.

²⁹ *Ibid* at para 437.

back to s 172(1), for success in proving a ground of review under s 6(2) of the PAJA is tantamount to showing that particular ‘conduct . . . is inconsistent with the Constitution’, which triggers s 172(1)(a).

When one looks at s 172(1)(a), one is immediately struck by the mandatory nature of the language. The court entertaining the constitutional matter has no choice but to declare the impugned legislation or conduct invalid. This position of course aligns with the doctrine of objective invalidity flowing from the supremacy clause of the Constitution, as explained in judgments such as *Ferreira v Levin*³⁰ and *Fose*.³¹ In the latter case Kriegler J stated that the supremacy clause ‘makes unconstitutional conduct a nullity, even before Courts have pronounced it so’ and that ‘it is not the declaration itself that renders the conduct unconstitutional. The declaration is merely descriptive of a pre-existing state of affairs.’³² The Court again confirmed this position recently in *Economic Freedom Fighters* when it held that

[d]eclaring law or conduct inconsistent with the Constitution and invalid is plainly an obligatory power vested in this court as borne out by the word ‘must’ [in s 172(1)]. Unlike the discretionary power to make a declaratory order in terms of s 38 of the Constitution, this Court has no choice but to make a declaratory order where s 172(1)(a) applies. Section 172(1)(a) impels this court, to pronounce on the inconsistency and invalidity of, in this case, the President’s conduct and that of the National Assembly. This we do routinely whenever any law or conduct is held to be inconsistent with the Constitution.³³

Within this constitutional framework it is not immediately apparent to us where the inherited common-law discretionary nature of the courts’ power in setting administrative action aside fits in. The core of our argument is that it is not axiomatic that the declaration of invalidity can mean anything but nullifying the administrative action. To say that the declaration of invalidity does not in effect amount to the setting aside of the administrative action seems to fly in the face of the doctrine of objective invalidity, the approach adopted in relation to remedies in judicial review of legislation and, arguably, s 172(1) itself. When one looks at the wording of s 172(1), it seems that para (b) assumes the real effect of the declaration of invalidity under para (a), that is to nullify the legislation or conduct. It is not altogether clear how else one would interpret the reference to ‘the retrospective effect of the declaration of invalidity’ in para (b) other than as suggesting that the declaration under para (a) results in nullifying the legislation or conduct.

Against this background we analyse the Constitutional Court’s approach in *Bengwenyama*,³⁴ *Allpay*³⁵ and *Kirland*.³⁶ In particular, we question the notion that the effect of nullifying the administrative action following review flows from the just and equitable order under s 172(1)(b) (and of course s 8 of the PAJA) rather

³⁰ *Ferreira v Levin NO and Others; Vryenhoek and Others v Powell NO and Others* [1995] ZACC 13, 1996 (1) SA 984 (CC), 1996 (1) BCLR 1 (CC).

³¹ *Fose v Minister of Safety and Security* [1997] ZACC 6, 1997 (3) SA 786 (CC), 1997 (7) BCLR 851 (CC) (‘Fose’).

³² *Ibid* at para 94.

³³ *Economic Freedom Fighters v Speaker of the National Assembly and Others* [2016] ZACC 11, 2016 (3) SA 580 (CC), 2016 (5) BCLR 618 (CC) at para 103 (footnotes omitted).

³⁴ *Bengwenyama Minerals (Pty) Ltd and Others v Genorah Resources (Pty) Ltd and Others* [2010] ZACC 26, 2011 (4) SA 113 (CC), 2011 (3) BCLR 229 (CC) (‘Bengwenyama’).

³⁵ Note 2 above.

³⁶ Note 1 above.

than the declaration of invalidity under s 172(1)(a). After analysing the cases, we return to the broad constitutional considerations outlined above as well as the text of s 172 to interrogate the fit between the discretionary nature of the remedy of setting aside at common law and the constitutional approach to remedies.

IV THE CONSTITUTIONAL COURT

The two 2014 cases of *Allpay 2* and *Kirland* built on the Constitutional Court's earlier pronouncement on this issue in *Bengwenyama Minerals (Pty) Ltd v Genorah Resources (Pty) Ltd*.³⁷ Our engagement with the Court's treatment of the discretionary nature of orders setting aside administrative action should thus start with a brief look at *Bengwenyama*.

A *Bengwenyama*

In judicial review proceedings the applicants in *Bengwenyama* challenged the grant of a prospecting right over their land to the first respondent by the Department of Mineral Resources. It was not in dispute that the decision to grant the prospecting right amounted to administrative action and that the PAJA applied to the decision. The review application was thus one brought in terms of the PAJA. The Court found the grant of the prospecting right reviewable on a number of PAJA grounds.

The discretionary nature of the setting-aside remedy was subsequently raised squarely. The respondent relied on a line of Supreme Court of Appeal judgments, *Oudekraal*,³⁸ *JFE Sapela Electronics*³⁹ and *Millennium Waste*,⁴⁰ to support an argument that the Court should not set the action aside despite the reviewable irregularities. The applicant argued in response that these judgments could not stand in light of s 172(1) of the Constitution, which makes a declaration of invalidity mandatory. Thus, the core of our puzzle was raised and very much in the terms outlined above: *Oudekraal* versus s 172(1).

The Court unfortunately does not provide much conceptual clarity on the issue. This is partly owing to some unfortunate formulations in its reasoning. The Court states that 'invalid administrative conduct must be declared unlawful'.⁴¹ Given that lawfulness is one of the constitutional requirements of just administrative action in s 33 and that invalidity is the label attached to conduct that does not comply with constitutional standards under s 172, one would have rather expected this statement to read: 'unlawful administrative conduct must be declared invalid'. The Court continues to refer to a declaration of unlawfulness, ostensibly referring to the declaration under s 172.⁴² The problem with this terminology is that it conflates the finding on the ground of review which, as we

³⁷ Note 34 above.

³⁸ Note 3 above.

³⁹ *Chairperson, Standing Tender Committee and Others v JFE Sapela Electronics (Pty) Ltd and Others* [2005] ZASCA 90, 2008 (2) SA 638 (SCA).

⁴⁰ *Millennium Waste Management (Pty) Ltd v Chairperson, Tender Board: Limpopo Province and Others* [2007] ZASCA 165, 2008 (2) SA 481 (SCA).

⁴¹ *Bengwenyama* (note 34 above) at paras 82 and 84.

⁴² *Ibid* at paras 82 and 85.

have argued above, amounts to a finding as to whether s 33 has been violated, with the declaration that must subsequently be made under s 172. In our view, a finding of unlawfulness is conceptually something different from a declaration of invalidity. The one expresses a view on whether the relevant conduct complies with the standard set for such conduct by s 33 of the Constitution, while the other pronounces on the result of that finding. In cases involving rights other than those in s 33, the distinction between these two pronouncements is typically made clear by the limitations analysis conducted in terms of s 36 of the Constitution. Keeping these two pronouncements distinct will aid us in identifying where there is judicial discretion in the review of administrative action under the Constitution and where there is none, as we shall argue below.

However, the most significant aspect of *Bengwenyama* is the distinction the Court draws between the declaration of invalidity and an order setting the administrative action aside. The Court holds that the order setting the administrative action aside forms part of ‘a further just and equitable remedy’ following the declaration of invalidity.⁴³ It is only the order setting aside the action, ie the further or consequential remedy, that is discretionary. This is made clear in the Court’s statement that ‘[t]he apparent rigour of declaring conduct in conflict with the Constitution and PAJA unlawful is ameliorated in both the Constitution and PAJA by providing for a just and equitable remedy *in its wake*’.⁴⁴

The Court also expressly rejects the possibility of grounding this approach in s 172(1)(b)(ii), which allows for the suspension of a declaration of invalidity.⁴⁵ The Court thus rejects the conceptualisation of a decision not to set aside the administrative action as a suspension of invalidity. In this respect the Court states that ‘[i]f the administrative action is declared unlawful, but all its consequences are not set aside, the practical effect of the order will be final, not merely a temporary suspension of invalidity’.⁴⁶ This statement is somewhat puzzling in terms of the Court’s own reasoning, in that it suggests an equation of setting aside and invalidity. It is ostensibly only the issue of timing that distinguishes the two positions, namely the declaration of unlawfulness without setting the action aside (fully) and the suspension of invalidity. The former is permanent while the latter is temporary, but the rest is the same.

Instead of relying on the ‘suspension’ conceptualisation in support of its approach, the Court relies on the general reference in s 172(1)(b) to a ‘just and equitable order’ as well as on s 8 of the PAJA, which echoes that provision.

B Allpay 2

Allpay 2 was the second instalment of the botched public tender for the rendering of services to pay social grants to beneficiaries on behalf of the South African Social Security Agency (SASSA). The case followed the familiar pattern in procurement disputes, with a disappointed bidder, Allpay, challenging the decision to award to

⁴³ Ibid at para 84.

⁴⁴ Ibid at para 85 (emphasis added).

⁴⁵ Ibid at para 82.

⁴⁶ Ibid.

the winning tenderer, Cash Paymaster Services (Pty) Ltd. In *Allpay 1*,⁴⁷ the first Constitutional Court judgment in this matter, the Court found the award decision reviewable on a number of PAJA grounds and thus held the decision to award the tender to Cash Paymaster ‘constitutionally invalid’.⁴⁸ However, the Court suspended that order ‘pending determination of a just and equitable remedy’ after further submissions by the parties. *Allpay 2* was the judgment that dealt with these further submissions and that formulated the remedy in this matter.

It is important to note the significance of the decision challenged in the *Allpay* saga, which played a key role in how the Court dealt with the matter and in particular the remedy. This was clearly not just another run-of-the-mill procurement case. Not surprisingly, the Court in *Allpay 1* framed the entire matter by highlighting the importance of the function at stake in this review. The opening paragraph of the judgment reads as follows:

For many people in this country the payment of social grants by the state provides the only hope of ever living in the material conditions that the Constitution’s values of dignity, freedom and equality promise. About 15 million people depend on the payment of these social grants. They are vulnerable people, living at the margins of affluence in our society.⁴⁹

The tender at issue was for the exclusive rendering of grant payment services to all beneficiaries across the country for a period of five years. By the time *Allpay 2* was heard, that service involved nearly 21 million people,⁵⁰ which is about 40 per cent of the total population.⁵¹ The subject matter of the service, social grants, is of course itself also recognised as a fundamental right under s 27 of the Constitution. From a public-function perspective this was accordingly a highly significant tender. From a commercial perspective it was no less significant. The winning bidder submitted that by the time the case was heard, it had already incurred capital expenditure of R1.3 billion under the contract.⁵² It was reported in the media that the tender was worth an estimated R10 billion.⁵³ The tender process itself took close to three years to complete and cost about R6 million.⁵⁴

Against the backdrop of the significance of the decision under review, in *Allpay 1* the Court was not prepared to rule on the appropriate remedy immediately on finding that the award decision was irregular, ie that SASSA had breached s 33 of

⁴⁷ *Allpay Consolidated Investment Holdings (Pty) Ltd and Others v Chief Executive Officer, South African Social Security Agency, and Others* [2013] ZACC 42, 2014 (1) SA 604 (CC), 2014 (1) BCLR 1 (CC) (*‘Allpay 1’*).

⁴⁸ *Ibid* at para 93.

⁴⁹ *Ibid* at para 1.

⁵⁰ *Allpay 2* (note 2 above) at para 14.

⁵¹ The 2011 census reported the total South African population at 51 770 560: Statistics South Africa *Census 2011*, Fact Sheet 1.

⁵² *Allpay 2* (note 2 above) at para 18.

⁵³ S Evans ‘Concourt Asked to “Turn its Back” on Unlawful Grants Tender’ Mail & Guardian (11 February 2014), available at <http://mg.co.za/article/2014-02-11-concourt-asked-to-turn-its-back-on-unlawful-grants-tender>; City Press ‘We Saved Government R3bn Per Year – Cash Paymaster Services’ news24 (11 May 2014), available at <http://www.news24.com/Archives/City-Press/We-saved-government-R3bn-per-year-Cash-Paymaster-Services-20150430>; South African Press Association ‘Social Development Welcomes Grant Ruling’ SABC News (18 April 2014), available at <http://www.sabc.co.za/news/a/b61ca50043ac3b99b5b9f7239b19c088/SocialundefinedDevelopmentundefinedwelcomesundefinedgrantundefinedruling-20141804>.

⁵⁴ *Allpay 2* (note 2 above) at para 12.

the Constitution in awarding the tender to Cash Paymaster. The Court referred to the fact that the system operating under the contract between SASSA and Cash Paymaster was running smoothly and that an abrupt setting aside of the tender award, and hence the contract, would potentially cause significant disruption.⁵⁵ The Court also noted that those likely to be most severely affected by such disruption were vulnerable members of society and in particular, children.⁵⁶ Accordingly, the Court called for additional submissions from the parties on a range of issues that would impact on the formulation of a remedy. These focused closely on what the likely implications might be should the tender award be set aside and how the interruption of the public function at stake here would be managed. In *Allpay 2* the Court dealt with these further submissions.

Already in *Allpay 1* the Court explicitly endorsed the approach in *Bengwenyama*,⁵⁷ stating, again in a somewhat odd formulation, that '[o]nce a ground of review under PAJA has been established there is no room for shying away from it. Section 172(1)(a) of the Constitution requires the decision to be declared unlawful.'⁵⁸ As in *Bengwenyama*, the Court conflated the finding of unlawfulness with the declaration of invalidity in this formulation. However, it held that the consequences of such a 'declaration of unlawfulness' are something distinct from the declaration itself, and must be dealt with in the subsequent just and equitable order. In *Allpay 1* the Court added that this approach, that is the distinction between constitutional invalidity and the remedy, is a constitutional development and not part of our common law.⁵⁹ This is the first indication that here we may not be dealing with the same common-law approach as espoused in *Oudekraal*.⁶⁰

Following this approach, the unanimous Court in *Allpay 2* viewed the matter at hand as what would be a just and equitable remedy *subsequent* to the declaration of invalidity in *Allpay 1*. In other words, the declaration of invalidity was not really at issue in the remedy part of this case, which is *Allpay 2*. This is a very clear indication that the Court views as two separate questions the declaration of invalidity and what the effect of that declaration should be. It is seemingly only in the second part that discretion has a role to play in formulating what the real effect will be of the declaration of invalidity, that is, in the formulation of the subsequent just and equitable remedy.

In *Allpay 2* the Court pointed to the 'corrective principle' governing the formulation of a just and equitable remedy, which flows from s 172 of the Constitution as well as '[l]ogic, general legal principle, the Constitution and the binding authority of this court' and 'the rule of law and principle of legality'.⁶¹ This corrective principle requires 'the consequences of invalidity to be corrected or reversed where they can no longer be prevented', at least as a default position.⁶² However, the Court went on to hold that there may be circumstances in which it

⁵⁵ *Allpay 1* (note 47 above) at para 96.

⁵⁶ *Ibid.*

⁵⁷ Note 34 above.

⁵⁸ *Allpay 1* (note 47 above) at para 25, and see also para 56.

⁵⁹ *Ibid* at para 26.

⁶⁰ Note 3 above.

⁶¹ *Allpay 2* (note 2 above) at paras 30–32.

⁶² *Ibid* at para 30.

would be just and equitable to deviate from the corrective principle, although it did not find it ‘desirable or even feasible’ to formulate a general statement of what those circumstances might be.⁶³ The Court simply held that it must be accepted ‘that the application of the corrective principle is not uniform’.⁶⁴

In assessing the particular circumstances of the case before it to determine whether there should be a deviation from the default position of the corrective principle, the Court pointed out that ‘a just and equitable remedy will not always lie in a simple choice between ordering correction and maintaining the existing position’.⁶⁵ This statement at least indicates clearly that a decision to maintain the administrative action, that is an order declaring that the administrative action at issue is of full force and effect, may be a just and equitable remedy following a declaration of invalidity.

C Kirland

The unanimity with which the Court dealt with these issues in *Bengwenyama* and *Allpay 2* was somewhat disturbed by the subsequent judgments in *Kirland*.⁶⁶ The facts of the *Kirland* case are rather peculiar and merit close attention.

In July 2006 and May 2007 Kirland applied to the Superintendent-General of Health in the Eastern Cape, Mr Boya, for permission to ‘establish’ two hospitals. Although Mr Boya decided not to approve the application, that decision was not communicated to Kirland because Mr Boya fell ill and an Acting Superintendent-General, Dr Diliza, took over. Dr Diliza subsequently informed Kirland by letter that its application had succeeded.⁶⁷ Kirland acted on the department’s approval in preparing to establish the hospitals. These events comprise all interaction between Kirland and the department until 20 June 2008, more than seven months later.⁶⁸ Thus, as far as Kirland was concerned it could proceed with the establishment of the hospitals.

Kirland’s situation changed upon receipt of a letter, dated 20 June 2008, from the Superintendent-General, Mr Boya, who had returned to work in the meantime. This letter informed Kirland that the approval had been ‘withdrawn’, the reason being that ‘Port Elizabeth [was] over serviced with private health facilities’.⁶⁹ Kirland appealed unsuccessfully to the MEC for Health in the Eastern Cape to have the withdrawal overturned, and subsequently initiated proceedings in the High Court. This was the complete picture as far as Kirland was concerned when the judicial proceedings began.

However, behind the scenes there was far more to the application than Kirland knew. At this stage nothing but the initial approval and subsequent withdrawal of the decision had been communicated to Kirland. In fact, before Mr Boya became ill he had followed an advisory committee’s recommendation to refuse Kirland’s

⁶³ Ibid at para 34.

⁶⁴ Ibid.

⁶⁵ Ibid at para 39.

⁶⁶ Note 1 above.

⁶⁷ *Kirland* (note 1 above) at para 16.

⁶⁸ Ibid at para 16.

⁶⁹ Ibid.

application. His decisions had been reduced to writing by 9 October 2007,⁷⁰ but were left unsigned and were not forwarded to Kirland.

Dr Diliza, the Acting Superintendent-General, was appointed in Mr Boya's absence. Upon learning of Mr Boya's decision to refuse the application, 'the MEC summoned the Acting Superintendent-General to her office on 23 October 2007'.⁷¹ According to Dr Diliza the MEC, claiming political pressure, instructed her to approve the application. Dr Diliza duly approved the application, under the impression that she was obliged to do so. This was denied by both Kirland and the MEC, but Kirland could not secure an affidavit from the MEC.

Even though Mr Boya returned to work in November 2007, it was only on 16 July 2008, seven months after his return, that he contacted Kirland for the first time. The wording of the letter is significant:

I refer to the above matter, more particularly the letter dated 23 October 2007 that the acting superintendent-general of this department addressed to you. In that letter you were informed that your applications for a licence in respect of the above hospitals had been approved. This approval is contrary to our view that the area is over supplied.

I regret to inform you that the Department has *withdrawn* the approval. I point out that on 9 October 2007 and after I had considered all applications, I decided to refuse the application because Port Elizabeth is over serviced with private health facilities.

I advise that you have a right to appeal in writing to the MEC for Health against my decision. That appeal must be lodged with the MEC within 60 days from the date of this letter and must set out the grounds of appeal.⁷²

Mr Boya thus created the impression that the department was entitled to withdraw the approval. (Only after Kirland had initiated judicial proceedings on the basis of the contested withdrawal were the irregularities mentioned for the first time in the state parties' answering affidavits, and this nearly two years after the withdrawal.) The letter set in motion the series of events culminating in the judgment of the Constitutional Court. The withdrawal was the catalyst. It was the decision to which Kirland responded, which Kirland appealed to the MEC and finally brought before the High Court; though that is not to say that the proceedings were limited to the contents of the letter and Kirland's response to it.

Against this background, the Constitutional Court divided mainly on the question whether the validity of Dr Diliza's approval of Kirland's application was an issue before the Court. The two dissenting judgments by Jafta J and Zondo J both held that the validity of that decision was indeed before the Court, whereas the majority judgment of Cameron J held that it was not. On the face of it, the main issue that the judges seemed to disagree on was *when* a court may or ought to entertain a review, which may seem like a question of procedure: can a court review and set aside an administrative action in the absence of a clear application to do so? However, underlying this difference on procedure was a lack of clarity as to exactly what the question of the validity of administrative action really means. It was in this context of a contested interpretation of what exactly was

⁷⁰ Ibid at para 8.

⁷¹ Ibid at para 9.

⁷² Ibid at para 16 as reproduced in the judgment of Jafta J (emphasis added).

before the Court, specifically relating to the validity question, that the first cracks seemed to appear in the unanimity of the Court as to how invalid administrative action should be dealt with under the Constitution.

Jafta J formulated the issue as ‘what should be the response from a court where serious maladministration and abuse of public power is established but there is no request for the review of the offending administrative action’.⁷³ He stated that this issue ‘goes to the heart of the role played by our courts in ensuring that public power is properly exercised within the bounds of the Constitution’.⁷⁴ Jafta J held that the validity of the approval was squarely before the Court and that the Court had to pronounce on it. He further found the approval to be invalid. In criticising the Supreme Court of Appeal’s handling of this matter, Jafta J stated:

It is apparent from the record that the Supreme Court of Appeal adopted an unduly narrow approach to the matter. In doing so, it left intact an administrative decision which that Court had found to be invalid – a decision which was made under circumstances described by that court as ‘a sorry tale of mishap, maladministration and at least two failures of moral courage’.⁷⁵

Although these statements may be read at least to question whether irregular administrative action of the type at issue here could be left intact, ie whether there is any discretion to uphold that action following the finding of invalidity, Jafta J steps back from such a position by confirming later in his judgment that a court has discretion to order a just and equitable remedy ‘[i]f the coming into effect of an order invalidating an administrative action would result in an injustice’.⁷⁶ He thus largely kept to the *Bengwenyama* approach, although with less fervour than the Court expressed in that matter or in *Allpay 1* and *Allpay 2*.

Jafta J placed much emphasis on the fact that the discretion does not apply to the declaration of invalidity,⁷⁷ and interestingly noted that the choice of a just and equitable remedy subsequent to the declaration of invalidity ‘does not include the reversal of what was done during the first stage’.⁷⁸

Zondo J broke ranks with the *Bengwenyama* orthodoxy much more explicitly in his dissenting judgment, where he stated:

The view that the applicants did not ask the high court to set aside the Acting S-G’s decision raises the question of what the difference is between asking a court to decide that a certain administrative decision is invalid and of no force and effect, and asking it to set such an administrative decision aside. In my view there is no difference in law. If a court decides that a certain administrative decision is invalid and of no force and effect, the position is as if that administrative decision was never taken in the first place. The same applies if a court sets aside an administrative decision. Therefore, where a litigant has asked a court to set aside an administrative decision or where he or she has asked a court to find the administrative decision to be invalid and of no force and effect, the result would be the same whichever of the two the court adopted.⁷⁹

⁷³ Ibid at para 28.

⁷⁴ Ibid.

⁷⁵ Ibid at para 38.

⁷⁶ Ibid at para 52.

⁷⁷ Ibid at paras 56 and 59–60.

⁷⁸ Ibid at para 61.

⁷⁹ Ibid at para 129.

In his majority judgment, however, Cameron J kept to the *Bengwenyama* approach and even explicitly confirmed *Oudekraal*. He held that where a party wishes to invalidate an administrative action effectively, that party should formally approach a court on review for an order declaring the administrative action invalid and setting it aside. Without a proper hearing on both issues, a court lacks jurisdiction to decide whether the particular action is effective. Following the approach in *Oudekraal*, the administrative action would stand.⁸⁰ One reason for this finding was in fact that the court should first consider the consequences of effective invalidity,⁸¹ that is of setting the administrative action aside, before it grants that order, very much as was done in *Allpay 1* and *Allpay 2*. It follows that those consequences must be properly placed before the court by all parties concerned before the court exercises its discretion whether to grant a remedy that may correct the administrative failure, or to sustain the action or follow some other route.

While following the *Bengwenyama* and *Allpay* approach to the issue of validity of administrative action, the majority in *Kirland* made at least two important contributions to the debate. First, the Court adopted terminology that is less confusing than that employed in the earlier judgments, as pointed out above.⁸² Cameron J used the terms ‘unjust administrative action’ or ‘irregular administrative action’ to refer to administrative action that has been found noncompliant with the Constitution, in particular s 33, via a ground of review.⁸³ This terminology is to be preferred for it avoids the confusion, at this stage of the inquiry, of using ‘lawful’, which refers to only one aspect of administrative justice, or ‘valid’ and ‘invalid’, which refers to the subsequent declaration. It thus avoids potentially conflating the finding of an infringement of a right and the subsequent declaration and remedy.

Secondly, Cameron J made the important point, albeit in a footnote, that ‘[t]here is no right to a perfect administration’.⁸⁴ This statement may provide a useful basis for exploring possible solutions to the puzzle at issue in this contribution. One could, for example, expand on this statement by arguing that the right to administrative justice does not equate to a right to administrative perfection. Consequently, even when administrative justice is being enforced, it may be that imperfect administrative action continues to be binding at times.

V CONCLUSION

The main difficulty with the Constitutional Court’s approach to the issue of validity of administrative action as expressed in *Bengwenyama*, *Allpay 2* and the majority judgment in *Kirland* is that the distinction made between the declaration of invalidity under s 172(1)(a) and the effect of nullifying the administrative action thus declared invalid in a *subsequent* order of setting aside is not wholly persuasive. As we have argued above, in terms of the constitutional framework it is not

⁸⁰ Ibid at para 106.

⁸¹ Ibid at para 86.

⁸² See text to note 41 above.

⁸³ *Kirland* (note 1 above) at paras 93, 96, 97.

⁸⁴ Ibid at para 88 fn 54.

axiomatic that the declaration of invalidity can mean anything but nullifying the administrative action.

It also does not seem to us to be an adequate answer to this puzzle to point to the familiar distinction between the common-law remedy of a declaratory order (later given a statutory foundation in the Supreme Court Act 59 of 1959) and an order setting aside administrative action. As the Court has itself noted, the declaration of invalidity under s 172 is 'a unique remedy created by the Constitution'.⁸⁵

To be sure, while we are not convinced of the Constitutional Court's reasoning regarding the proper conceptualisation of the discretionary nature of the setting-aside remedy under the Constitution, it does seem highly desirable to have flexibility in managing the effects of invalidating administrative action on review. Many of the problems we experience in judicial review of administrative action can arguably be ascribed to a blunt approach to remedies, mostly the knee-jerk setting aside of reviewable administrative action.⁸⁶

What we are questioning is not the desirability of the discretionary nature of remedies in judicial review of administrative action, in particular relating to the invalidation of such action, but rather how that discretion ought to be accommodated within our constitutional setting. It seems to us that the common-law position as expressed for instance in *Oudekraal* cannot continue, but that the two-stage approach of the Constitutional Court as expressed in *Bengwenyama*, *Allpay* and *Kirland* is not ideal either. Perhaps a better approach would be to focus on the stage of constitutional adjudication preceding the application of s 172. At least the possibility should be explored of incorporating the desired discretion into the assessment of the impugned conduct against s 33 or of developing a limitation clause analysis in administrative-law review. Exploring each of these options in any detail goes beyond the scope and goal of this contribution and would require in-depth analysis of administrative-law review within the approach to and structure of fundamental rights adjudication that has developed under the Constitution. We accordingly offer only brief remarks on these options.

Hoexter has long argued persuasively for variability in the application of administrative-justice standards.⁸⁷ She argues that variability 'allows the courts to be more generous about the application of administrative justice and to vary its precise content according to the circumstances'.⁸⁸ The court can thus adjust the content of what administrative justice entails in a given context depending on the circumstances. The basic notion of flexibility in the content of particular administrative justice standards has also long been recognised in the context of procedural fairness and has been explicitly recognised in the constitutional era,

⁸⁵ *Islamic Unity Convention v Independent Broadcasting Authority and Others* [2002] ZACC 3, 2002 (4) SA 294 (CC), 2002 (5) BCLR 433 (CC) at para 10.

⁸⁶ See G Quinot 'Towards Effective Judicial Review of State Commercial Activity' 2009 *Journal of South African Law* 436; G Quinot 'Worse than Losing a Government Tender: Winning It' (2008) 19 *Stellenbosch Law Review* 101.

⁸⁷ C Hoexter 'The Future of Judicial Review in South African Administrative Law' (2000) 117 *South African Law Journal* 484, 502–505; C Hoexter *Administrative Law in South Africa* (2nd Edition, 2012) 222–223.

⁸⁸ Hoexter 2012 (note 87 above) at 222.

both statutorily⁸⁹ and judicially.⁹⁰ It seems to us that part of the considerations informing a variable application of administrative justice standards could be the need to maintain the relevant administrative action. The type of substantive reasoning applied by the Constitutional Court in *Joseph* in respect of the content of procedural fairness under the PAJA is arguably an example of such an approach.⁹¹ The court reasoned that s 3(2)(a) of the PAJA must be read ‘as an empowering provision that allows courts to exercise a discretion in enforcing the minimum procedural fairness requirements under s 3(2)(b)’.⁹² One finds in this reasoning the type of discretion in the application of administrative justice standards that could, in suitable circumstances, serve the same purpose as the common-law remedial discretion in respect of setting aside administrative action.

The second option would be to develop a limitations analysis approach to the adjudication of breaches of administrative-justice rights in line with the adjudication of all other fundamental rights. Up to now s 36 of the Constitution has played virtually no role in judicial review of administrative action in terms of the Constitution. The most obvious reason is undoubtedly the s 36 requirement of a ‘law of general application’ for the limitation of rights. Without delving here into a detailed analysis of this requirement and how it has been interpreted to date, there is arguably room for the view that administrative action properly taken under an empowering provision that qualifies as a ‘law of general application’ will satisfy the requirements of s 36. In such a case a court could thus find that a ground of review exists, meaning that the impugned administrative action infringes s 33 of the Constitution, but that the infringement amounts to a justifiable limitation under s 36 and is thus not constitutionally invalid. Section 172(1) is accordingly not triggered. On this approach there would still be a weighing up of the competing interests that are typically taken into account when a court decides whether to exercise its remedial discretion to set aside irregular administrative action – but the weighing would effectively be done in terms of the structured proportionality analysis of s 36.

Developing these approaches to judicial review of administrative action further would bring administrative-law review much closer to judicial review generally under the Constitution. In this development the distinction drawn by Cameron J between perfect administrative action and just administrative action may be a particularly useful conceptual tool.

⁸⁹ Section 3(2)(a) of the PAJA states that ‘[a] fair administrative procedure depends on the circumstances of each case’.

⁹⁰ See, eg, *Joseph and Others v City of Johannesburg and Others* [2009] ZACC 30, 2010 (4) SA 55 (CC), 2010 (3) BCLR 212 (CC) (*Joseph*) at para 58.

⁹¹ *Ibid.*

⁹² *Ibid* at para 59.

The Ebb and Flow of the Application of the Principle of Subsidiarity – Critical Reflections on *Motau* and *My Vote Counts*

Melanie Murcott*
Werner van der Westhuizen†

I INTRODUCTION

Under the current constitutional dispensation the judiciary is not only constitutionally authorised, but also constitutionally obliged, to oversee exercises of public power; including the conduct of the executive.¹ It does so through judicial review. In judicial review proceedings, courts must follow a principled and justified approach to choosing the appropriate standards on a possible ‘continuum of constitutional accountability’ against which impugned exercises of public power should be measured. This is what is demanded by the separation-of-powers doctrine: courts ought not to invoke legal norms formalistically or arbitrarily when reviewing public power.²

In determining where on the continuum of constitutional accountability an exercise of public power should lie, we argue that subsidiarity theory plays a valuable role – particularly in the context of administrative law, where several sources of law compete with one another for application. At one end of the continuum lie the most foundational norms of accountability, such as the rule of law, a founding value in the Constitution of the Republic of South Africa, 1996.³ These foundational norms are the more general legal norms that root and create the context in which the more detailed and indirect constitutional norms

* Senior Lecturer, Department of Public Law, Faculty of Law, University of Pretoria.

† Candidate Attorney, Bowmans.

The authors are very grateful for the input received at the seventh Constitutional Court Review conference in December 2015, and for the invaluable input of the reviewers and editors. Any errors in the work are our own.

¹ We support the approach to the separation of powers as articulated in D Mosenke ‘Oliver Schreiner Memorial Lecture: Separation of Powers, Democratic Ethos and Judicial Function’ (2008) 24 *South African Journal on Human Rights* 341, 349 where it is recognised that the courts ‘not only have the right to intervene in order to prevent the violation of the Constitution, they also have the duty to do so’.

² We agree with R Stacey ‘Justifiability as the Animating Vision of Administrative Law’ (2007) 22 *SA Public Law* 79, 103 who states that ‘actions on the part of review courts that ignore the proper justification for judicial control of the executive might weaken the legitimacy of the judicial review mechanism’.

³ Section 1(c) of the Constitution.

find their application.⁴ Legality, a constitutional principle inherent in the rule of law, lies at this end of the continuum of accountability. Next on the continuum lie norms of accountability found in the Bill of Rights, such as those contained in the rights to just administrative action in s 33 of the Constitution. Section 33 is aimed at realising the rule of law in relation to exercises of public power that amount to administrative action. In the middle of the continuum lie the indirect constitutional norms aimed at achieving accountability, such as those contained in the Promotion of Administrative Justice Act 3 of 2000 (PAJA). The statute gives effect to and fleshes out the content of the rights to just administrative action: it is intended to provide the most immediate source and guidelines for judicial review of conduct that amounts to administrative action.⁵ Still further along the continuum lie even more specific empowering provisions in other statutes or in subordinate legislation. These empowering provisions are the most specific norms that set out standards of accountability demanded of a functionary in a particular situation, and that are appropriate to that specific exercise of power.

Legality is intended as a basis on which to review only those exercises of public power that do not amount to administrative action as ‘a backstop or safety net ... when the PAJA [is] not of application’.⁶ In other words, placing exercises of public power at their correct point on the continuum of constitutional accountability in judicial review entails applying PAJA and/or other more specific norms when they are applicable, and applying legality only if and to the extent that more specific norms are not applicable to hold the power to account.

Given the intended roles of PAJA and legality respectively, one could be forgiven for thinking, more than fifteen years after the enactment of PAJA, that it would be trite that exercises of public power amounting to ‘administrative action’ as defined in PAJA ought to be reviewed on PAJA grounds. This is all the more so in view of the endorsement of subsidiarity theory by the Constitutional Court as early as 1995.⁷ Put simply, subsidiarity requires that adjudication of substantive issues be determined with reference to more particular constitutional norms, rather than more general constitutional norms.⁸ As we illustrate in this article, applying subsidiarity theory in the administrative-law context entails that, in judicial review proceedings, the applicability of PAJA ought to be determined

⁴ We see specific norms as emerging indirectly from the Constitution, and general norms as emerging directly from the Constitution.

⁵ *Zondi v MEC for Traditional and Local Government Affairs and Others* [2004] ZACC 19, 2005 (3) SA 589 (CC), 2005 (4) BCLR 347 (CC) at para 101; *Bato Star Fishing (Pty) Ltd v Minister of Environmental Affairs and Others* [2004] ZACC 15, 2004 (4) SA 490 (CC), 2004 (7) BCLR 687 (CC) (‘*Bato Star*’) at para 25, where it was confirmed that the cause of action for the review of administrative action is now rooted in PAJA. See also J Klaaren & G Penfold ‘Just Administrative Action’ in S Woolman & M Bishop (eds) *Constitutional Law of South Africa* (2nd Edition, OS, 2008) Chapter 63 at 63–64.

⁶ C Hoexter ‘The Enforcement of an Official Promise: Form, Substance and the Constitutional Court’ (2015) 132 *South African Law Journal* 207 (‘Enforcement’), 219.

⁷ In *S v Mhlungu and Others* 1995 ZACC 4, 1995 (3) SA 867 (CC), 1995 (7) BCLR 793 (CC) (‘*Mhlungu*’).

⁸ Ibid at para 59. See further L du Plessis ‘Subsidiarity: What’s in the Name for Constitutional Interpretation and Adjudication?’ (2006) 17 *Stellenbosch Law Review* 207; and see also J de Visser ‘Institutional Subsidiarity in the South African Constitution’ (2010) 1 *Stellenbosch Law Review* 90. In this article we use the term ‘subsidiarity theory’ in the sense of ‘adjudicative subsidiarity theory’ described by Du Plessis. See also *Bato Star* (note 5 above). A more nuanced account of subsidiarity is provided below.

before resort is had to the principle of legality.⁹ However, flouting subsidiarity theory, the courts, including the Constitutional Court, often invoke legality as a basis for reviewing exercises of public power without deciding whether PAJA is applicable, even though it might be.¹⁰

In this article we reflect on the principle of subsidiarity in theory with reference to its content and origins, and also on its practical application, particularly in the context of administrative law. We focus on two recent decisions of the Constitutional Court: *Minister of Defence and Military Veterans v Motau*¹¹ and *My Vote Counts v Speaker of the National Assembly*.¹² We argue that in *Motau*, an administrative-law case, the Constitutional Court applied subsidiarity theory correctly, and gave some indication of how the courts ought to approach the threshold questions that plague them when reviewing exercises of public power. However, we argue that *Motau* did not go far enough. It did not provide the clarity that courts reviewing exercises of public power require as to how and why subsidiarity theory ought to be adopted. We argue that *My Vote Counts*, a case concerning the relationship between the Promotion of Access to Information Act 2 of 2000 (PAIA) and s 32 of the Constitution, in which the Constitutional Court resoundingly endorsed the principle of subsidiarity in constitutional adjudication, offers guidance on the invocation of the principle in other contexts, including administrative-law disputes.

Before commenting on *Motau* and *My Vote Counts*, we outline the jurisprudential context within which they were decided. This context illustrates both the inconsistent application of the principle of subsidiarity as well as the need for a coherent and principled approach to determining whether and when legality or PAJA should be invoked as the basis for reviewing exercises of public power. We do not focus on the courts' avoidance of PAJA as a result of its cumbersome definition and procedural requirements, a subject that has been dealt with extensively elsewhere.¹³ Rather, we seek to explain why subsidiarity theory

⁹ *Minister of Health and Another NO v New Clicks South Africa (Pty) Ltd and Others* [2005] ZACC 14, 2006 (2) SA 311 (CC), 2006 (1) BCLR 1 (CC) ('*New Clicks*') at para 97. See also Hoexter (note 6 above) at 221.

¹⁰ We will demonstrate that whereas in *New Clicks* (note 9 above) and *Bato Star* (note 5 above) the Court explicitly and zealously invoked the principle of subsidiarity, in *Albutt v Centre for the Study of Violence and Reconciliation and Others* [2013] ZACC 10, 2010 (3) SA 291 (CC), 2010 (5) BCLR 391 (CC) ('*Albutt*') and *KwaZulu-Natal Joint Liaison Committee v MEC, Department of Education, KwaZulu Natal and Others* [2013] ZACC 10, 2013 (4) SA 262 (CC), 2013 (6) BCLR 615 (CC) ('*KZN JLC*') the Court failed to do so. We focus our analysis on *Albutt* as the first case in which the Constitutional Court abandoned subsidiarity in the context of administrative law, and on *KZN JLC* as a missed opportunity to apply subsidiarity theory correctly in the post-*Motau* context.

¹¹ *Minister of Defence and Military Veterans v Motau and Others* [2014] ZACC 18, 2014 (5) SA 69 (CC) ('*Motau*'). As Hoexter points out in 'Enforcement' (note 6 above) at 223 (references omitted): '[I]n *Motau* the Constitutional Court noted unobtrusively but explicitly that "[t]he correct order of enquiry is to consider, first, whether PAJA applies, and only if it does not, what is demanded by general constitutional principles such as the rule of law". This seems to contradict the *Albutt* approach and, with respect, to make far better sense.'

¹² *My Vote Counts NPC v Speaker of the National Assembly and Others* [2015] ZACC 31, 2016 (1) SA 132 (CC) ('*My Vote Counts*').

¹³ See, eg, Hoexter 'Enforcement' (note 6 above) at 220–221; C Hoexter 'The Rule of Law and the Principle of Legality in Administrative Law Today' in M Carnelley & S Hoorntje (eds) *Law, Order and Liberty: Essays in Honour of Tony Matthews* (2011) 55 ('Rule of Law'); M Murcott 'Procedural Fairness as a Component of Legality: Is a Reconciliation between *Albutt* and *Masetlha* possible?' (2013) 130 *South African Law Journal* 260 ('Procedural Fairness'), 266–270.

offers a basis for solving the threshold questions of whether and when PAJA as opposed to legality ought to be invoked in judicial review proceedings. Our central contention is that courts ought to answer the threshold questions of administrative law in a coherent and principled manner that balances the values of constitutional supremacy and constitutional democracy; and that subsidiarity theory offers a basis for doing so.¹⁴

II SUBSIDIARITY THEORY

A Content and Origin

Subsidiarity theory is richly explored in German legal theory.¹⁵ It is, however, not uniquely German, nor is it primarily a legal term; rather, subsidiarity is generally considered a Roman-Catholic social doctrine.¹⁶ Put simply, subsidiarity recognises the implicit hierarchies in communities and proposes a bottom-to-top approach: the lower level of the hierarchy should in principle exhaust its capacity to contribute in a particular context before the higher level intervenes, either by taking over or providing assistance where the lower level has reached its limit.¹⁷ Viewed negatively, the higher level plays only a supportive – subsidiary – role. Viewed positively, the higher level creates the context and circumstances within which the lower level operates.¹⁸ Van Wyk argues that, in principle, subsidiarity could apply to any hierarchical relationship and in various contexts as part of so-called ‘unwritten constitutional law’.¹⁹ Du Plessis notes that subsidiarity is not inevitably confined to hierarchically arranged norms.²⁰ It can also refer to the law’s interpretative preference for the application of a particular norm over another applicable but subsidiary norm that is required to ‘step down’ in a particular context.²¹

Subsidiarity as endorsed in South African law encapsulates the notion that adjudication of substantive issues should (subject to certain provisos that we discuss below) be determined with reference to more particular, indirect constitutional norms applicable, rather than more general, direct constitutional

¹⁴ Below we endorse aspects of the views of AJ van der Walt ‘Normative Pluralism and Anarchy: Reflections on the 2007 Term’ (2008) 1 *Constitutional Court Review* 77 (‘Normative Pluralism’) to the effect that subsidiarity theory is ‘an angle of approach’ that allows courts to adopt a coherent approach to solving complex questions about which sources of law to apply in constitutional adjudication.

¹⁵ D van Wyk ‘Subsidiariteit as Waarde wat die Oop en Demokratiese Suid-Afrikaanse Gemeenskap ten Grondslag Lê’ in G Carpenter (ed) *Suprema Lex: Essays on the Constitution Presented to Marinus Wiechers* (1998) 251, 257–259 (Recognises the fundamental role that subsidiarity plays in human rights in a *Rechtsstaat*, where tension clearly exists between the individual’s self-determination and dignity and the state’s coercive powers and authority. Van Wyk’s essay is one of the few works in South African law exclusively devoted to an investigation into subsidiarity theory. While it focuses on subsidiarity theory as applied to the structure of government under the Constitution, the work is relevant here as it explains the roots of the theory).

¹⁶ Ibid at 254 (Notes that the theory is not peculiar to the Roman-Catholic tradition and can also be found in Protestantism).

¹⁷ Ibid.

¹⁸ Ibid at 243–255.

¹⁹ Ibid at 255.

²⁰ Du Plessis (note 8 above) at 208.

²¹ Ibid.

norms applicable.²² General, direct constitutional norms are consequently not overused, since more specific legal precepts solve the legal question.²³ As early as 1995 the Constitutional Court endorsed subsidiarity theory in the context of constitutional adjudication (adjudicative subsidiarity theory) in *S v Mhlungu*.²⁴ The Court set out the guiding principle of subsidiarity that courts should prefer an outcome, if that is justified in the particular case, which does not involve the direct determination of a constitutional issue.²⁵ In *Zantsi v Council of State, Ciskei* the court further invoked subsidiarity to justify the methodology of developing the law incrementally.²⁶

Van der Walt views the Court's development of subsidiarity theory as an 'angle of approach' that incorporates a balancing of the values of constitutional rights (or supremacy) and democracy.²⁷ In cases involving the alleged infringement of a constitutional right, he identifies two subsidiarity principles, with provisos, followed by the courts at the threshold stage to determine the source of law in terms of which the constitutional analysis should commence. The first subsidiarity principle guides the choice between the Constitution and legislation as potential sources of law in relation to an alleged infringement of a constitutional right. The second subsidiarity principle deals with the choice between the common law and legislation as potential sources of law in relation to an alleged infringement of a constitutional right. Both principles, according to Van der Walt, affirm normative pluralism: the recognition of the contingent and contested nature of legal norms and values; but by offering a basis for choosing between the norms and values, the principles avoid adopting an 'anything goes' approach, which would amount to 'normative anarchy'.²⁸

The first subsidiarity principle, based on *SANDU v Minister of Defence*,²⁹ requires a litigant to rely on legislation when enforcing a constitutional right rather than circumventing the legislation in favour of direct application of a constitutional provision; with the proviso that the constitutional provision may be invoked

²² Ibid at 215–223.

²³ Ibid at 215.

²⁴ *Mhlungu* (note 7 above) at para 59; see also Du Plessis (note 8 above) at 209, where the author describes the adjudicative subsidiarity in *Mhlungu* as a 'reading strategy'.

²⁵ *Mhlungu* (note 7 above) at para 59. See the discussion of Van der Walt below, where we address the circumstances fully.

²⁶ *Zantsi v Council of State, Ciskei and Others* [1995] ZACC 9, 1995 (4) SA 615 (CC).

²⁷ Van der Walt 'Normative Pluralism' (note 14 above) at 99. A comprehensive account of the emergence of a normative and legal pluralism falls outside the scope of this article. We reflect on Van der Walt's account only as one possible narrative to situate the development of subsidiarity theory in the Constitutional Court.

²⁸ Van der Walt 'Normative Pluralism' (note 14 above) at 78–80. For a critical analysis of Van der Walt's account of subsidiarity, see KE Klare 'Legal Subsidiarity and Constitutional Rights: A Reply to AJ Van der Walt' (2008) 1 *Constitutional Court Review* 129. For the purposes of this article, we align ourselves with Van der Walt's response to Klare in AJ van der Walt *Property and Constitution* (2012) 36 ('*Property*'). See also Woolman's critique of the indirect application of the Bill of Rights in S Woolman 'The Amazing, Vanishing Bill of Rights' (2007) 124 *South African Law Journal* 762.

²⁹ *South African National Defence Union v Minister of Defence and Others* [2007] ZACC 10, 2007 (5) SA 400 (CC), 2007 (8) BCLR 863 (CC) ('*SANDU*').

where such legislation is challenged for inconsistency with the supreme law.³⁰ This principle derives from s 39(2) of the Constitution, which requires the courts to give effect to legislation enacted by the legislature pursuant to, and within the limits of, its constitutional responsibilities.³¹ The principle confirms the constitutional obligation of the elected legislature to promote the spirit, purport and object of the Bill of Rights. Consequently it affirms the value of democracy by requiring that legislation be given effect by the judiciary as demanded by the separation of powers doctrine.³² The proviso, on the other hand, confirms the value of the supremacy of the Constitution and its fundamental rights by subjecting the democratic will of the legislature to the boundaries of the Constitution that the judiciary must uphold. In this way, subsidiarity in constitutional adjudication can be employed to find a *via media* between formalistic fundamentalism and an anything-goes approach, through candidly weighing and balancing the values of constitutional supremacy and democracy in each case.³³

Van der Walt's second subsidiarity principle, derived from *Bato Star*, prohibits a litigant from indirectly enforcing and protecting a constitutional right against infringement by means of the common law where legislation, which has been intended to codify the common law, gives effect to the right in question.³⁴ The right is protected by applying the legislation in conformity with the Constitution and invoking the common law only to interpret the legislation, in line with s 39(2) of the Constitution.³⁵ This principle is subject to the proviso that the common law may be invoked to protect the right only where the legislation does not give effect to the right (or simply does not cater for it), as long as the common law is not inconsistent with applicable constitutional rights or the legislative scheme, and then only where the common law cannot be developed in order to bring it in line with the Constitution.³⁶

These two principles of subsidiarity support Du Plessis's view that, even though the Constitution enjoys normative superiority to other legal sources, its

³⁰ Van der Walt *Property* (note 28 above) at 36. This proposition was similarly endorsed by the Constitutional Court in *Mazibuko and Others v City of Johannesburg and Others* [2009] ZACC 28, 2010 (4) SA 1 (CC), 2013 (11) BCLR 1297 (CC) at para 73 (Emphasised that litigants' causes of action should be derived from legislation that gives effect to a constitutional right instead of from the constitutional right itself, or should alternatively challenge the constitutional validity of the legislation).

³¹ Van der Walt 'Normative Pluralism' (note 14 above) at 100.

³² Ibid at 103.

³³ Van der Walt 'Normative Pluralism' (note 14 above) at 103. In this way Van der Walt contends subsidiarity affirms normative pluralism. This approach is supported by Du Plessis (note 8 above) at 227. He argues that because norms not contained in the Constitution itself, such as those in legislation, are required to be fundamentally aimed at giving effect to the values of the Constitution and promoting the rights in the Bill of Rights, the enforcement of these norms should be preferred where they appropriately give effect to those objectives. In constitutional adjudication where the issue is not identified as exclusively determinable by direct constitutional application, the fundamental rights and values should be enforced through the ordinary legal principles instead of applying the wider and more flexible constitutional norms directly.

³⁴ Van der Walt *Property* (note 28 above) at 36.

³⁵ Van der Walt 'Normative Pluralism' (note 14 above) at 103.

³⁶ Van der Walt *Property* (note 28 above) at 36. Van der Walt 'Normative Pluralism' at 103 (Contends that the second subsidiarity principle and proviso affirm normative pluralism and reject the development of a system of law alternative to the Constitution).

existence does not displace ordinary legal principles.³⁷ Where a constitutional norm is consistent in its function and objectives with more particular norms not contained in the Constitution itself ('non-constitutional norms'), subsidiarity theory supports the application of the latter.³⁸ The provisos recognise the fact that, where the more particular norm is inconsistent with the subsidiary general norm, the legal authority of the constitutional norm cannot be conferred on the non-constitutional norm. Accordingly, the impugned non-constitutional norm must be invalidated unless the norm can be developed in the light of the constitutional norm to assume its lawful position in the subsidiarity relationship.³⁹ This approach harmonises the myriad non-constitutional legal principles with the spirit, purport and objects of the Bill of Rights and the Constitution,⁴⁰ which in turn supports the notion of a single system of law as endorsed by the Constitutional Court in *Pharmaceutical Manufacturers*.⁴¹

B Applying Subsidiarity Theory to Administrative Law

Arguing that a subsidiarity approach should be adopted to the threshold determination of whether PAJA or the principle of legality is to be invoked as the source of law in judicial review proceedings, Hoexter states that in line with the precedent in *New Clicks*, 'PAJA must be applied where it is applicable'.⁴² Read in isolation, this statement presents a *prima facie* logical fallacy of circular reasoning: it states its own conclusion. It is for this reason that the question may be posed: when is PAJA applicable? This seems to be a pre-threshold question to determine whether PAJA and the principle of legality are indeed competing for application in a particular dispute. We contend that

an 'administrative-law case' should be conceived broadly, as one in which the conduct in question *might* (but does not necessarily) amount to 'administrative action' in terms of s 1 of the PAJA, alternatively s 33 of the Constitution.⁴³

This conception of a broader 'administrative-law case' is premised on the proposition that the relationship between PAJA and the principle of legality is one of subsidiarity and one that recognises the continuum of constitutional accountability. While the main contenders as sources of law during judicial review proceedings are the principle of legality and PAJA, they are in fact the particular precepts of two broader constitutional norms of accountability. Legality is an instance of the rule of law, the foundational value enshrined in s 1(c) of the Constitution. PAJA, on the other hand, is the constitutional legislation mandated by s 33(3) to give effect to the constitutional rights to just administrative action,

³⁷ Du Plessis (note 8 above) at 226. According to Du Plessis, 'adjudicative subsidiarity [provides opportunities] to de-absolutise the power of the Constitution'.

³⁸ Ibid.

³⁹ Ibid.

⁴⁰ Ibid.

⁴¹ *Pharmaceutical Manufacturers Association of South Africa and Another; In re Ex Parte President of the Republic of South Africa and Others* [2000] ZACC 1, 2000 (2) SA 674 (CC), 2000 (3) BCLR 241 (CC) ('*Pharmaceutical Manufacturers*') at para 45.

⁴² C Hoexter *Administrative Law in South Africa* (2nd Edition, 2012) ('*Administrative Law*') at 134.

⁴³ Murcott 'Procedural Fairness' (note 13 above) at 268.

and is the conduit for the indirect application of the s 33 rights. Therefore, at a fundamental level, the primary (direct) norms of accountability that are competing for application are the rule of law as a founding value of the Constitution, and the rights to just administrative action enshrined in s 33. The constitutional analysis must accordingly take into account these underlying norms to construct a complete picture of the relationship between PAJA and the principle of legality at the threshold stage of the analysis.⁴⁴

The subsidiary norm in this relationship is the rule of law, whereas the more specific norm is s 33 of the Constitution, and the most specific norm is PAJA, giving effect to s 33.⁴⁵ The textual setting of the rule of law is key to understanding the nature of the norm and why it indeed constitutes the subsidiary norm, and its consequent place on our continuum of accountability in relation to exercises of public power. Since the rule of law is enshrined as a founding value of the Constitution, its normative relevance differs from that of a substantive right in Chapter 2 of the Constitution. Even though conduct inconsistent with the founding values is considered unlawful, the values are used primarily to inform the interpretation of the Constitution and of legislation and to drive the development of the common law, thus infusing the normative edifice of democracy.⁴⁶ For this reason, the constitutional design envisages that the abstract founding values on the one end of the continuum of accountability should shape the development of the other provisions of the Constitution and the ordinary law lying towards the other end of the continuum.⁴⁷ The values must be applied directly only where the other particular legal precepts have been exhausted. This approach to adjudication thus requires the more particular precept to be applied first, before resort is had to the more general value.

Whereas the rule of law generally requires the state to exercise its powers in accordance with the law,⁴⁸ the principle of legality has developed as a justiciable instance of the rule of law that imposes standards for the exercise of all public power.⁴⁹ The principle of legality is thus a general norm of constitutional law, and in adjudication this character must be taken into account at the threshold stage when considering its possible application to the case at hand.

The rights to administrative justice in s 33 of the Constitution constitute justiciable fundamental rights. This section of the Constitution has constitutionalised administrative law⁵⁰ and imposes particular standards of

⁴⁴ For us the continuum of accountability is a helpful way to construct the complete picture as it reflects the range of norms of accountability from the most general (subsidiary) norms to the most specific norms.

⁴⁵ As we discuss above, there will, of course, also be empowering provisions imposing standards of accountability in respect of particular exercises of public power, which represent even more specific norms of accountability.

⁴⁶ I Currie & J de Waal *The Bill of Rights Handbook* (6th Edition, 2013) at 7–8.

⁴⁷ Ibid at 8. See also *Minister of Home Affairs v National Institute for Crime Prevention and the Re-Integration of Offenders (NICRO) and Others* [2004] ZACC 10, 2005 (3) SA 280 (CC), 2004 (5) BCLR 445 (CC) ('NICRO') at para 21, discussed in Currie & De Waal (note 46 above) at 8 fn 33.

⁴⁸ Currie & De Waal (note 46 above) at 10.

⁴⁹ FI Michelman 'The Rule of Law, Legality and the Supremacy of the Constitution' in S Woolman & M Bishop (eds) *Constitutional Law of South Africa* (2nd Edition, OS, 2010) Chapter 11 at 11–14.

⁵⁰ *Pharmaceutical Manufacturers* (note 41 above).

administrative justice where conduct amounts to administrative action, a particular subset of exercises of public power.⁵¹ Therefore, whereas the rule of law and the principle of legality generally require the state to rule under law, s 33 further develops this fundamental norm by fleshing out the legal standards required where conduct amounts to administrative action.

PAJA, in turn, is the product of the constitutional mandate in s 33(3) and is the codifying and reformative legislation that gives effect to s 33(1) and (2). The legislation is the conduit for the indirect enforcement of the rights to just administrative action in s 33 and prescribes the specific standards of administrative law. This is why litigants are not at liberty to invoke s 33 directly and must source their cause of action in PAJA when exercises of public power amount to administrative action. Its standards apply to all ‘administrative action’ and in turn supplement any particular rules of administrative law found in legislation or regulations. Even though the principle of legality has become known as a parallel system of administrative law, it does not constitute any particular rules of administrative law that must be applied before the supplementary application of PAJA. The principle of legality is in fact a broader norm of constitutional law that regulates all exercises of public power, and PAJA is therefore a more specific norm of this constitutional value.

For these reasons the relationship between the rule of law and the s 33 rights is one of subsidiarity. The rule of law, and in turn the principle of legality, constitutes the more general and subsidiary norm, whereas s 33 and, in particular, PAJA, contain the primary rules of administrative law. PAJA, even though it enjoys the status of constitutional legislation,⁵² is one step removed from the Constitution, since it is a statute. The principle of legality is, however, a constitutional norm. Therefore, invocation of the principle of legality amounts to a direct application of a constitutional norm, whereas the application of PAJA constitutes the application of a norm one step removed from the Constitution, and thus an indirect application of s 33 of the Constitution, and more broadly, the rule of law.

This relationship can be read in terms of the positive and negative conceptions of subsidiarity theory. The rule of law as a founding value not only supports, but creates the context within which s 33 and PAJA operate. The supportive role of the principle of legality is especially evident in its function as a safety net where conduct does not amount to administrative action and the more particular norms have been exhausted. The principle of legality accordingly ensures that exercises of public power do not escape the continuum of accountability created by the Constitution. The subsidiary relationship between PAJA and the principle of legality explains when PAJA is applicable in a particular case and why our conception of a broader ‘administrative-law case’ should be supported. The rule of law, in turn, is the broader constitutional value. Whenever the potential application of the demands of administrative justice is in issue, unless specialised legislation is applicable and PAJA itself assumes a supplementary role,⁵³ the point

⁵¹ Currie & De Waal (note 46 above) at 13.

⁵² Ibid at 12–13.

⁵³ I Currie *The Promotion of Administrative Justice Act: A Commentary* (2nd Edition, 2007) (‘*Commentary*’) at para 1.4.

of departure ought to be PAJA. If it were not for this inquiry, ‘only supernatural perspicacity on the part of the court would allow it to know in advance that there is no reason to “reach” the PAJA’.⁵⁴ As Hoexter puts it:

The first question in any administrative-law case ought surely to be whether the most specific and most detailed norm, the PAJA, is applicable, and not whether the problem is capable of being solved by the rule of law, a far more general and abstract constitutional doctrine.⁵⁵

Since the principle of legality is subsidiary to PAJA in the broader administrative-law case, the *SANDU* subsidiarity principle and proviso should be applied. The principle requires a litigant to source her cause of action in legislation when enforcing a constitutional right. As the court confirmed in *New Clicks*, the litigant must rely directly on PAJA.⁵⁶ The court must first apply the threshold provisions of PAJA in order to determine whether the conduct in question amounts to administrative action.⁵⁷ This means that the litigant cannot rely directly on s 33, the more general constitutional norm from which PAJA is derived and to which the legislation gives effect. Furthermore, following the reasoning in *NICRO*, a litigant ought not to rely directly on the rule of law before determining whether the more particular norms (such as s 33, and by virtue of s 33(3), the PAJA) are applicable.⁵⁸ Accordingly, the Constitution is appropriately applied indirectly through the ordinary legal provisions. In the language of subsidiarity, the more particular norms must first be exhausted before resort is had to the more general norms. What subsidiarity teaches us in this process is that the supremacy of the Constitution must be upheld, and since PAJA is constitutional legislation, a purposive interpretation of PAJA must be applied.⁵⁹

The proviso to the *SANDU* principle is that the right to which the legislation gives effect can be invoked directly when attacking the constitutional validity of the legislation. This proviso is confirmed by *New Clicks*,⁶⁰ and it is accordingly open to the litigant to invoke s 33 directly when the constitutional validity of PAJA is called into question. Where the conduct in question does not amount to administrative action and PAJA is construed and deemed consistent with s 33 of the Constitution, the more particular norm of administrative justice has clearly been exhausted. Since s 33 and PAJA flesh out the more general norm of the rule of law, the threshold requirement that conduct must amount to administrative action serves to limit the exercises of power that are subject to the more rigorous precepts of administrative justice.⁶¹ Therefore, subsidiarity permits the more general norm to intervene and take over where the particular norms have run out. It is following this approach that the court in *New Clicks* endorsed the

⁵⁴ Hoexter ‘Enforcement’ (note 6 above) at 222.

⁵⁵ Hoexter *Administrative Law* (note 42 above) at 134.

⁵⁶ *New Clicks* (note 9 above) at paras 96–97.

⁵⁷ Murcott ‘Procedural Fairness’ (note 13 above) at 269.

⁵⁸ *NICRO* (note 47 above) at paras 20–25; Currie & De Waal (note 46 above) at 8 & 13.

⁵⁹ Currie *Commentary* (note 53 above) at 2.1.

⁶⁰ *New Clicks* (note 9 above) at para 97.

⁶¹ Hoexter *Administrative Law* (note 42 above) at 120. Although legality has begun to mirror PAJA’s standards, this is, in part, due to the incorrect application of subsidiarity. Moreover, legality’s standards of accountability remain more flexible and thus more uncertain than those of PAJA.

safety-net function of the principle of legality.⁶² When the particular conduct in question does not amount to administrative action (for example, when it amounts to executive action), the principle of legality serves its subsidiary function by providing a safety net that imposes standards of constitutional accountability further along the continuum of accountability:⁶³ the more general norm of the rule of law takes over where the more specific norm of administrative justice has run out.

The invocation of the principle of legality can also be seen as filling an intentional gap in PAJA. As Van der Walt argues, such a gap indeed exists in PAJA with the effect that the review of conduct that does not amount to administrative action takes place by direct reliance on the constitutional principle of legality (as a safety net).⁶⁴ Regrettably, post-*New Clicks*, legality has not always played this safety-net function. The judicial appreciation for subsidiarity theory has ebbed. In the next part of this article we comment on the ebb of subsidiarity theory with reference to *Albutt*⁶⁵ and *KZN JLC*.⁶⁶ We then consider the revival of subsidiarity theory in the Constitutional Court's jurisprudence in 2015, first in *Motau*,⁶⁷ an administrative-law case, and later in *My Vote Counts*,⁶⁸ a case concerning the relationship between the constitutional right to access to information and the statute giving effect to that right.

III SUBSIDIARITY IN PRACTICE

A The Ebb of Subsidiarity Theory: *Albutt* and *KZN JLC*

Although the Constitutional Court confirmed as early as 2004 that the cause of action for the review of administrative action is now rooted in PAJA,⁶⁹ Hoexter points out that

[f]rom the PAJA's inception there was widespread resistance to the statute, or more accurately to its overly elaborate definition of administrative action, and the courts frequently indulged applicants who preferred not to engage with it. Sometimes the court noted that the principle of legality was capable of resolving the matter at hand, while in other instances the PAJA was simply ignored without explanation. This pattern of avoidance culminated in the extraordinary judgment of the Constitutional Court in *Albutt*.⁷⁰

In *Albutt* the Constitutional Court found conduct of the President to be irrational through application of the principle of legality.⁷¹ The court criticised the High Court for considering and applying the provisions of PAJA, and found it

⁶² *New Clicks* (note 9 above) at para 97.

⁶³ Murcott 'Procedural Fairness' (note 13 above) at 270.

⁶⁴ Van der Walt 'Normative Pluralism' (note 14 above) at 106.

⁶⁵ *Albutt* (note 10 above).

⁶⁶ *KZN JLC* (note 10 above).

⁶⁷ *Motau* (note 11 above).

⁶⁸ *My Vote Counts* (note 12 above).

⁶⁹ *Bato Star* (note 5 above) at para 25.

⁷⁰ Hoexter 'Enforcement' (note 6 above) at 221.

⁷¹ *Albutt* (note 10 above) at para 74.

unnecessary to apply PAJA itself.⁷² In considering the possible application of PAJA the court first expressed the opinion that, in light of its prior jurisprudence, the conduct of the President was ‘unlikely’ to amount to administrative action.⁷³ This important preliminary conclusion was not reached by applying PAJA’s definition of administrative action, however. The court pointed out that, should the power amount to administrative action under PAJA, a number of ‘complex questions’ would have to be answered which it found ‘unnecessary’ to address.⁷⁴ Rejecting an inquiry into the applicability of PAJA by invoking judicial minimalism or avoidance, Ngcobo J found that:

Sound judicial policy requires us to decide only that which is demanded by the facts of the case and is necessary for its proper disposal.... At times it may be tempting, as in the present case, to go beyond that which is strictly necessary for a proper disposition of the case. Judicial wisdom requires us to resist the temptation.... There may well be cases, and they are very rare, when it may be necessary to decide an ancillary issue in the public interest. This is not such a case.⁷⁵

The *Albutt* approach to determining administrative law’s threshold questions has, on the one hand, been followed in countless subsequent administrative-law cases,⁷⁶ and on the other, been severely criticised amongst legal scholars precisely because the approach flies in the face of subsidiarity theory.⁷⁷ The *Albutt* approach flouts that theory because it countenances the invocation of a general norm (legality) even where a norm of greater specificity (PAJA) might be applicable.

The consequences of the rejection of subsidiarity in *Albutt* are threefold. First, the court’s minimalist approach resulted in a failure to adopt a principled and justified approach to choosing the appropriate standards on the possible continuum of accountability against which the President’s conduct ought to have been tested. The court’s choice of legality rather than PAJA as a basis for reviewing the President’s conduct appears arbitrary. Secondly, the *Albutt* approach undermined the principle of democracy and the separation of powers. This is because the legislature has, in PAJA, articulated the standards of natural justice required when the President’s exercises of public power do amount to administrative action. By failing to consider whether those standards were applicable, the court disregarded the legitimate role of the legislature in setting those standards. Moreover, the court’s failure to justify, properly, on substantive grounds, the basis upon which it would not apply PAJA, undermined the separation of powers. Finally, the court detracted from constitutional supremacy, in that PAJA is the constitutionally mandated legislation that gives effect to s 33 of the Constitution which courts *must* invoke when reviewing exercises of public power that amount to administrative action.

⁷² Ibid at para 76.

⁷³ Ibid at para 80.

⁷⁴ Ibid at para 81.

⁷⁵ Ibid at para 82.

⁷⁶ See, eg, the cases discussed in D Brand & M Murcott ‘Administrative Law’ 2013 *Annual Survey of South African Law* 61, 61–69.

⁷⁷ See, eg, Hoexter *Administrative Law* (note 42 above) at 131, Murcott ‘Procedural Fairness’ (note 13 above) at 269–270 and Hoexter ‘Enforcement’ (note 6 above) at 221–223.

In *Valuline CC v Minister of Labour*, the Minister of Labour exercised her power in terms of the Labour Relations Act 66 of 1995 to extend a collective agreement to non-signatories in the clothing manufacturing industry.⁷⁸ The litigants challenged the Minister's decision on the bases of both the principle of legality and PAJA.⁷⁹ Koen J, however, found it 'irrelevant' to determine whether the conduct in question amounted to administrative action, and consequently whether PAJA was applicable, since the conduct fell to be reviewed under the principle of legality.⁸⁰ In following the *Albutt* approach, Koen J's selection of legality as a basis to review the Minister's conduct disregarded the legitimate role of the legislature in enacting PAJA so as to give effect to s 33 of the Constitution.

Given cases such as *Valuline*, one would think that the Constitutional Court would seek to clarify the relationship between PAJA and legality.⁸¹ Instead of doing so, in *KZN JLC* the court further muddled the judicial waters. In *KZN JLC* the Constitutional Court invoked rationality as an aspect of the rule of law (legality) as the basis upon which to review the revocation of an official promise by the provincial Department of Education to pay subsidies to independent schools.⁸² As Hoexter has argued, in *KZN JLC* public law was 'permitted to come to the rescue notwithstanding the applicant's non-reliance on the PAJA'.⁸³

During 2008, the Department promised that independent schools in the province would be paid subsidies in 2009–2010. In 2009, however, the department sought to revoke its official promise and reduce the subsidies.⁸⁴ The schools sought to enforce the department's promise, but in doing so, expressly disavowed reliance upon public law.⁸⁵ Accordingly, the Constitutional Court was not, on an explicit reading of the pleadings before it, confronted with a choice of reviewing the department's conduct on the basis of either PAJA or legality. When asked by the *amicus* (who was not party to the proceedings until the matter was brought to the Constitutional Court) to hold the department's conduct to account in terms of PAJA, the court declined to do so on the basis that the record of the department's decision was not before it as it would have been had a review application been

⁷⁸ *Valuline CC and Others v Minister of Labour and Others* [2013] ZAKZPHC 9, 2013 (4) SA 326 (KZP) ('*Valuline*'). Another pre-*Motau* example of the application of the *Albutt* approach is *Southern African Litigation Centre and Another v National Director of Public Prosecutions and Others* [2012] ZAGPPHC 61, 2012 (10) BCLR 1089 (GNP) at para 18, noted in Hoexter 'Enforcement' (note 6 above) at 222.

⁷⁹ *Valuline* (note 78 above) at para 7.

⁸⁰ *Ibid.*

⁸¹ See Brand & Murcott (note 76 above) at 61–69 where they illustrate a need for clarity on whether and when to apply PAJA or legality in cases concerning the review of exercises of public power that might amount to administrative action.

⁸² *KZN JLC* (note 10 above) at para 71. See further Hoexter 'Enforcement' (note 6 above); M Murcott 'A Future for the Doctrine of Substantive Legitimate Expectation? The Implications of *KwaZulu-Natal Joint Liaison Committee v MEC for Education, KwaZulu-Natal*' (2015) 18 *Potchefstroom Electronic Law Journal* ('Substantive Legitimate Expectation') 3133.

⁸³ Hoexter 'Enforcement' (note 6 above) at 219.

⁸⁴ *KZN JLC* (note 10 above) at paras 2–8.

⁸⁵ *Ibid* at para 31.

brought in terms of PAJA.⁸⁶ Instead the Court offered the schools a ‘public-law lifeline’ through the application of broad, general constitutional-law norms.⁸⁷

The court’s reasoning in relation to the rejection of PAJA as a basis for review is thin, for the absence of a record is not (despite the court’s contentions to the contrary) a principled reason on which to select the basis for holding the conduct of the executive branch to account.⁸⁸ Rather than apply PAJA, the court invoked broad principles of public law as the bases upon which to hold the department to account and compel the department to fulfil its publicly promulgated promise to pay the schools. These principles were reliance, accountability and rationality.⁸⁹ While we do not fault the court for invoking *a* basis upon which to hold government to account, we do find it problematic that the court selected broad principles under the rule of law without canvassing *substantive reasons* why PAJA ought or ought not to apply.

The court held that sufficient facts had been pleaded to found a remedy in public law,⁹⁰ but other than the absence of a record, did not make it clear why that public-law remedy could not arise from PAJA rather than broad principles. The effect of failing to determine whether the more specific statutory provisions under PAJA were applicable was again to disregard the legitimate role of the legislature in enacting those provisions. Thus the court undermined the principle of democracy and the separation of powers. In doing so, the court resorted to legality without adequately justifying why PAJA could be circumvented – even though the conduct bore ‘all the hallmarks of administrative action’ under PAJA.⁹¹ In this way the *KZN JLC* Court implicitly endorsed the *Albutt* approach in that it saw no need to justify, on substantive grounds, why PAJA could be overlooked.

Although *KZN JLC* is somewhat anomalous in the manner it was pleaded and argued, like *Albutt* it illustrates a failure by the Constitutional Court to appreciate the need to adopt a coherent and principled approach in selecting a basis to review exercises of public power and hold government to account when conduct might amount to administrative action in terms of PAJA. It is in this jurisprudential context that the tide was due for a change when *Motau* was decided towards the end of 2014.

B The Flow of Subsidiarity Theory: *Motau* and My Vote Counts

Motau concerned the review of a decision by the Minister of Defence and Military Veterans (the Minister) to remove the Chair, General Motau, and the Deputy Chair, Ms Mokoena, (the respondents) from the Board of Directors of the Armaments Corporation of South Africa (SOC) Ltd (Armcor).⁹² Armcor is a wholly state-owned entity regulated by the Armaments Corporation of South Africa Limited

⁸⁶ Ibid at para 32.

⁸⁷ Hoexter ‘Enforcement’ (note 6 above) at 219.

⁸⁸ See ibid and at 224, where Hoexter points out that the court was ‘patently uninhibited by the absence of the record when it reasoned on broader and more abstract constitutional lines’. See also Murcott ‘Substantive Legitimate Expectation’ (note 82 above) at 3142–3143 and 3146.

⁸⁹ *KZN JLC* (note 10 above) at para 63.

⁹⁰ Ibid at para 70.

⁹¹ Murcott ‘Substantive Legitimate Expectation’ (note 82 above) at 3135.

⁹² *Motau* (note 11 above) at para 1.

Act 51 of 2003 (Armscor Act) over which the state exercises ownership control through the Minister. The affairs of Armscor are managed and controlled by the board, comprising nine executive and two non-executive members. Section 7(1) and (2) of the Armscor Act empower the Minister to appoint the non-executive members of the board and to designate two of these members as the Chair and Deputy Chair of the board. The Minister is in turn empowered by s 8(c) of the Armscor Act to remove the non-executive members from office. It provides that '[a] member of the Board must vacate office if his or her services are terminated by the Minister on good cause shown'.⁹³

Acting pursuant to these provisions, the Minister terminated the services of the respondents on various grounds,⁹⁴ arguing that her decision did not involve a 'legal matter' but rather a 'political matter ... informed by [her] experience'.⁹⁵ Dissatisfied with the Minister's decision, the respondents approached the High Court and sought to have her decision set aside as unlawful, unconstitutional and invalid.⁹⁶ The High Court agreed with the respondents' contentions. It found that the Minister's decision amounted to administrative rather than executive action and ruled that the Minister had, amongst other things, made an error of law and failed to afford the respondents due procedural fairness.⁹⁷ The court accordingly set aside the decision for failing to comply with the standards of accountability envisaged by the PAJA. Apart from ruling that the decision of the Minister was unlawful, the court also held that the Minister had failed to show the good cause for removal required by s 8(c) of the Armscor Act. By failing to identify the particular lapses for which the respondents could be held liable, she had unfairly singled out the respondents in light of the board's collective responsibility for managing the affairs of Armscor.⁹⁸

The Constitutional Court's approach to reviewing the Minister's decision is significant. In the first paragraph of the judgment, Khampepe J made it clear that the case turned on accountability and that the Court was required to determine the standards against which it ought to hold the Minister to account in exercising her powers of oversight in respect of the board of Armscor.⁹⁹ The Court thus made it clear at the threshold stage of the constitutional analysis that it was vitally important to determine which legal norms on the continuum of accountability should be invoked in order to hold the Minister to account.

The Court in *Motau* adopted subsidiarity theory in its approach to determining the appropriate standards of accountability by first ascertaining whether the conduct amounted to administrative action in order to determine whether the PAJA was applicable. The Court's reliance on subsidiarity theory is evident in the following passage found in its threshold analysis of the legal problem:

⁹³ Section 8(c) of the Armscor Act.

⁹⁴ See *Motau* (note 11 above) at paras 9–16.

⁹⁵ *Ibid* at para 15.

⁹⁶ *Ibid* at para 16: *Motau and Another v Minister of Defence and Military Veterans and Another* (GNP case no 51258/13).

⁹⁷ See *Motau* (note 11 above) at paras 18–19.

⁹⁸ *Ibid* at para 20.

⁹⁹ *Ibid* at para 1.

Does the Minister's decision amount to administrative or executive action? Answering this question is important. If it amounts to administrative action, it is subject to a higher level of scrutiny in terms of PAJA. If it is executive action, it is subject to the less exacting constraints imposed by the principle of legality.¹⁰⁰

Importantly, the Court relied on the subsidiarity principle identified in *SANDU* as the basis for choosing the appropriate standard of accountability.¹⁰¹ In the footnote to the quotation above, the Court quoted O'Regan J's statement in *SANDU* that 'a litigant who seeks to assert [a constitutional right] should in the first place base his or her case on any legislation enacted to regulate the right, not [the Constitution]'.¹⁰²

Having answered the threshold question on the basis of subsidiarity theory, the Court subsequently embarked on the administrative action analysis in order to determine whether PAJA, identified as the potential standard of accountability, was in fact applicable. The Court concluded that the decision of the Minister amounted to executive rather than administrative action. This finding was primarily based on the fact that the Minister's decision involved the formulation of policy in respect of Armscor in the broad sense rather than the implementation of policy in the narrow sense.¹⁰³ Having satisfied itself that the standards of accountability encompassed by PAJA were not applicable, the Court invoked the rationality standard imposed by the principle of legality as a safety net, and found that the Minister's conduct had indeed been rational.¹⁰⁴

In answering the question whether there were any procedural restraints on the Minister's exercise of her power to remove the respondents from the board, the Court first considered whether the accountability standards contained in s 71(1) and (2) of the Companies Act 71 of 2008 were applicable. This provision imposes procedural requirements on the removal of a director from the board by a company's shareholders.¹⁰⁵ The Court found that s 8(c) of the Armscor Act had to be read together with s 71(1) and (2) of the Companies Act, and concluded that the Minister acted unlawfully in failing to afford the respondents the procedural fairness dictated by the legislation.¹⁰⁶ In its *obiter* remarks, the Court inquired whether in the absence of s 71 the principle of legality would impose procedural restraints on the exercise of the Minister's powers. After briefly examining the contentious cases of *Masetlba*¹⁰⁷ and *Albutt*,¹⁰⁸ which dealt with the question of whether legality encompasses a procedural component, the Court concluded, consistent with subsidiarity theory, that this question 'does not, in the light of the applicability of the Companies Act, need to be decided here'.¹⁰⁹

¹⁰⁰ Ibid at para 27.

¹⁰¹ Ibid at para 27 fn 28.

¹⁰² *SANDU* (note 29 above) at para 52.

¹⁰³ *Motau* (note 11 above) at paras 44 and 51.

¹⁰⁴ Ibid at paras 69–71.

¹⁰⁵ Ibid at paras 72–73.

¹⁰⁶ Ibid at paras 74–80.

¹⁰⁷ *Masetlba v President of the Republic of South Africa and Another* [2007] ZACC 20, 2008 (1) SA 566 (CC), 2008 (1) BCLR 1 (CC) ('*Masetlba*').

¹⁰⁸ *Albutt* (note 10 above).

¹⁰⁹ *Motau* (note 11 above) at para 83.

The court in *Motau* therefore approached not only the relationship between PAJA and the principle of legality on the basis of subsidiarity theory, but also the relationship between the Companies Act and the principle of legality. As Hoexter notes, even though natural justice or procedural fairness is undoubtedly part of the rule of law, it has been uncertain whether it is a requirement of the principle of legality.¹¹⁰ In contrast with the *Albutt* approach, the court's subsidiarity approach to the question whether the Minister was constrained by procedural fairness standards in *Motau* is a sound exercise in judicious avoidance.¹¹¹ By recognising that the more specific norm, the Companies Act, fleshes out the procedural fairness requirements of the more general norm, the rule of law, the Court avoided the overuse of a direct constitutional norm and the further extension of the principle of legality. It appropriately upheld the principle of democracy, affirming that the legislature's expression of the standards of procedural fairness required in the particular case of shareholders removing directors from a board of a company should be exhausted before resort is had to the constitutional principle of legality. The rule of law creates the context in which the detailed legislative norms can more appropriately provide nuanced standards of accountability in particular circumstances. The legislature has greater capacity to determine what natural justice demands in specialised situations such as the removal of a director in company law.

The approach of the Court in *Motau* should be lauded for its correct application of subsidiarity theory in answering the threshold question of determining the standards on the continuum of accountability against which the impugned exercise of public power ought to be measured. However, the Constitutional Court could have gone further than adopting subsidiarity theory implicitly and merely referring to the authority for its approach in a footnote. As was argued above, the jurisprudence reveals that the courts have failed to follow a consistent approach to determining which standards of accountability ought to be applied in judicial review proceedings, resulting not only in confusing but also in arbitrary decision-making. The Constitutional Court thus missed a valuable opportunity to revisit its reasoning in *Albutt* and expound further on a principled application of the subsidiarity theory that underpinned its reasoning.

However, soon afterwards the Constitutional Court handed down judgment in a case that will likely become the *locus classicus* of constitutional subsidiarity theory, and that can perhaps make up for the Court's failures in *Motau*. In *My Vote Counts* the Court resoundingly endorsed subsidiarity theory in the context of the relationship between the right to access to information in s 32 of the Constitution and the legislation giving effect to that right, PAIA.¹¹² Although *My Vote Counts*

¹¹⁰ Hoexter *Administrative Law* (note 42 above) at 418. See further Hoexter's discussion of the contentious cases of *Albutt* (note 10 above) and *Masetlha* (note 107 above) at 418–420 regarding the question whether the principle of legality demands procedural fairness.

¹¹¹ See I Currie 'Judicious Avoidance' (1999) 15 *South African Journal on Human Rights* 147. The Court appropriately left open the question when legality might demand the application of procedural-fairness standards, as it relied instead on the more specific rules contained in the Companies Act. The court thus placed the exercise of the power at the appropriate point on the continuum of accountability by relying on the most specific norms available.

¹¹² *My Vote Counts* (note 12 above).

was not an administrative-law case, it should illustrate to courts and litigants that it offers a sound approach for resolving the threshold questions of administrative law.

My Vote Counts concerned a non-profit organisation that promoted accountability, transparency and inclusiveness in the South African political system. Concerned with the prominent and largely unregulated role afforded to political parties under the Constitution, the applicant, *inter alia*, campaigned for law reform regarding the funding of political parties. In an attempt to achieve such reform, it resorted to the courts. Relying on the constitutional right of access to information, the applicant applied directly to the Constitutional Court, seeking an order to compel the legislature to enact legislation requiring ‘systematic and proactive disclosure of private funding of political parties’.¹¹³

The nub of the applicant’s case depended on a direct application of s 32 of the Constitution.¹¹⁴ This section is divided into two parts. The first part guarantees the right to access any information held by the state and private persons. Whereas the right to access state information is expressed in unconditional language, the right to access private information vests only if such information is required for the exercise or protection of any rights. The second part mandates the legislature, in peremptory terms similar to s 33(3) of the Constitution, to enact national legislation to give effect to this right. Accordingly, PAIA, the constitutional legislation that a party must now ordinarily rely on to enforce the right of access to information, was enacted.¹¹⁵

Critically, instead of sourcing its cause of action in PAIA, the applicant in *My Vote Counts* relied directly on s 32. According to the applicant, voters had the right to access information on private funding of political parties because such information is necessary for the exercise and protection of the right to vote guaranteed in s 19(3) of the Constitution.¹¹⁶ Furthermore, the legislature was required to enact national legislation to give effect to the right of access to information in s 32(1). However, because neither PAIA nor any other legislative provision required the systematic disclosure of information relating to party political funding, the legislature had failed to comply with its constitutional obligation fully to give effect to the right, and should therefore be ordered to enact legislation regulating the continuous disclosure of such funding.¹¹⁷

¹¹³ Ibid at para 8.

¹¹⁴ Section 32 of the Constitution provides: ‘(1) Everyone has the right of access to— (a) any information held by the state; and (b) any information that is held by another person and that is required for the exercise or protection of any rights. (2) National legislation must be enacted to give effect to this right, and may provide for reasonable measures to alleviate the administrative and financial burden on the state.’

¹¹⁵ Hoexter *Administrative Law* (note 42 above) at 97. This rule is based on the principle of adjudicative subsidiarity.

¹¹⁶ *My Vote Counts* (note 12 above) at para 19. Political parties had been characterised as private bodies in *Institute for Democracy in South Africa and Others v African National Congress and Others* [2005] ZAWCHC 30, 2005 (5) SA 39 (C). In order to satisfy the condition in s 32(1)(b), the applicant therefore argued that the information on private funding was necessary for the enforcement and protection of the right to vote.

¹¹⁷ *My Vote Counts* (note 12 above) at para 19.

In response, counsel for the legislature argued that the applicant's approach flouted the principle of subsidiarity. Applying this principle, the applicant was not entitled to circumvent PAIA and rely directly on s 32.¹¹⁸ Instead, because PAIA is the more specific norm that gives effect to s 32, the applicant ought to have enforced its right through the legislation. On the other hand, had the applicant contended that PAIA failed to give full effect to the constitutional right of access to information, the correct approach according to the principles of jurisdiction would have been to challenge the constitutionality of PAIA 'frontally' in the High Court.¹¹⁹ As a result, the applicant was caught in what the Court termed a 'logical trap': whether or not PAIA fully gave effect to its right of access to information, the application had to be dismissed.¹²⁰

The Court split 7-4 against the applicant. The crucial difference between the majority and the minority judgments was whether the principle of subsidiarity applied to the facts of the case: the majority dismissed the application on the basis that it did. The judgment commences with the dissent in which Cameron J provided a lucid historical account of the application and importance of the principle of subsidiarity in South African jurisprudence.¹²¹ Significantly, the majority of the Court concurred in Cameron J's exposition,¹²² thereby providing definitive full-court confirmation of the prominence of subsidiarity as a principle of South African constitutional law. After considering a number of the various applications of the broader concept of legal subsidiarity, the minority focused on the form of the principle of subsidiarity most applicable to the facts, namely adjudicative subsidiarity as expounded in this article. According to Cameron J, this form of subsidiarity has frequently been invoked to 'describe the principle that limits the way in which litigants may invoke the Constitution to secure enforcement of a right'.¹²³ Although all law and conduct derive their validity from the supreme Constitution, the Constitution's influence is ordinarily exerted indirectly through legislation and the common law.¹²⁴ Accordingly, the Court in *SANDU* confirmed the adjudicative subsidiarity principle and proviso that, where legislation is enacted to give effect to a constitutional right, a litigant may not invoke the Constitution 'without first relying on, or attacking the constitutionality of' such legislation.¹²⁵ Cameron J then cited a number of cases in which the courts have unequivocally applied the *SANDU* principle and proviso, including the *New Clicks* ruling that a litigant may not circumvent the PAJA in favour of the common law or the s 33 rights.¹²⁶

¹¹⁸ Ibid at paras 44–46.

¹¹⁹ Ibid at paras 122, 178 and 193. Section 172(2)(a) of the Constitution provides: 'The Supreme Court of Appeal, a High Court or a court of similar status may make an order concerning the constitutional validity of an Act of Parliament, a Provincial Act or any conduct of the President, but an order of constitutional invalidity has no force unless it is confirmed by the Constitutional Court.'

¹²⁰ *My Vote Counts* (note 12 above) at para 45.

¹²¹ Ibid at paras 47–64.

¹²² Ibid at para 121.

¹²³ Ibid at para 50.

¹²⁴ Ibid at para 52.

¹²⁵ Ibid at para 53.

¹²⁶ *New Clicks* (note 9 above) at para 96.

Moreover, both the minority and majority judgments made it clear that the principle of subsidiarity is not merely a formalistic tool or instrumentality invoked when determining threshold issues in constitutional adjudication. Summarising the minority's reasoning, the majority held that there are at least three important reasons for a subsidiarity approach to the application of the Constitution:

First, allowing a litigant to rely directly on a fundamental right contained in the Constitution, rather than on legislation enacted in terms of the Constitution to give effect to that right, 'would defeat the purpose of the Constitution in requiring the right to be given effect by means of national legislation'. Second, comity between the arms of government enjoins courts to respect the efforts of other arms of government in fulfilling constitutional rights. Third, 'allowing reliance directly on constitutional rights, in defiance of their statutory embodiment, would encourage the development of "two parallel systems" of law'.¹²⁷

These reasons emphasise the important value of subsidiarity: it promotes a principled method for determining when the Constitution should be applied directly or indirectly. Subsidiarity permits a substantive approach to applying legal norms, encouraging the adjudicator to recognise the values that underlie them. As Van der Walt argues, the *SANDU* principle and proviso emphasise that when legislation has been enacted giving effect to a right, the values of constitutional supremacy and democracy are subject to each other: to ignore legislation in favour of the direct application of a constitutional right where it exists 'would be to fail to recognise the important task conferred upon the Legislature by the Constitution to respect, protect, promote and fulfill the rights in the Bill of Rights'.¹²⁸ In turn, as evidenced by the majority judgment in *My Vote Counts*, this approach forced the Court candidly to reflect on the separation-of-powers implications of judicial review, thereby promoting principled and substantive judicial reasoning.

On the facts, the minority ultimately found that the principle of subsidiarity was not applicable. According to the Cameron J, the principle of subsidiarity did not find application owing to the manner in which the applicant pleaded its case. Rather than challenging the constitutional validity of PAIA, the applicant contended that the legislature had failed in its s 32(2) obligation to give effect to the right of access to information.¹²⁹ Because of the restricted ambit of PAIA, which fails to include a mechanism for systematic access to information on the funding of political parties, the legislature had failed to give proper effect to the right. Consequently, the applicant did not seek to circumvent PAIA in favour of the constitutional right and in violation of subsidiarity. Instead, 'the applicant confronted [PAIA] head-on, and invoked the Constitution only as a means to show that PAIA's reach falls short of fulfilling the obligations of Parliament under s 32(2)'.¹³⁰ As explained below, this should be viewed as an example of what Van der Walt calls a gap in legislation: where such a gap exists on a particular

¹²⁷ *My Vote Counts* (note 12 above) at para 160.

¹²⁸ *SANDU* (note 29 above) at para 52, quoted with approval in *My Vote Counts* (note 12 above) at para 163.

¹²⁹ *My Vote Counts* (note 12 above) at para 67.

¹³⁰ *Ibid* at para 74.

point, the subsidiarity principles ought not to apply and the litigant may resort directly to either the Constitution or the common law.¹³¹

The majority of the court, however, found this reasoning unconvincing based on the following approach. First, in terms of the case law and on an interpretation of the preamble, long title and objects of the Act, PAIA was the legislation intended to give effect to s 32 of the Constitution, notwithstanding the fact that there might be other legislative provisions also providing for access to information.¹³² Secondly, the crux of the applicant's objection was the manner in which the legislature had exercised its legislative powers. It argued that the legislature had failed in its duty to give effect to the right of access to information in that PAIA did not provide for the systematic disclosure of political parties' private funding. However, the majority held that in the absence of a constitutional challenge to PAIA, the separation-of-powers doctrine prevented a litigant from asking the court to prescribe to the legislature how to fulfil its legislative function.¹³³ In the circumstances, the applicant's argument amounted to nothing more than stating that PAIA was unconstitutional to the extent that it failed to give full effect to the right of access to information.¹³⁴ Accordingly, on an application of the proviso to the *SANDU* principle, subsidiarity required the litigant to challenge the constitutionality of PAIA frontally for its alleged shortcomings and deficiencies. Subsidiarity thus sounded the death-knell for the applicant's case. As the majority explained:

[W]e cannot bring ourselves to hold that there has been non-compliance with a constitutional obligation in circumstances where the shortcomings complained of by the applicant – and amplified by the minority judgment – may well prove to be constitutionally compliant. The issue is not whether they are indeed compliant. Whether they are, is something that may be tested properly in what we have tagged a frontal challenge. Therein lies the jurisprudential value of the principle of subsidiarity.

...

On the procedure resorted to by the applicant and the approach adopted by the minority judgment, the usual procedural hoops in a frontal challenge that invokes inconsistency with a right in the Bill of Rights are bypassed. It may well be that Parliament might have been able to demonstrate that what shortcomings there may be are justified in terms of s 36(1) of the Constitution. How do we then reach a conclusion that Parliament has failed to comply with a constitutional obligation? Or, do we simply say, quite plainly, Parliament could never have been able to show justification? How can we say that when – as we seek to demonstrate below – that was not a case that Parliament had to meet and, therefore, not an issue before us? That cannot be so.¹³⁵

In our view, the minority was correct in ruling that the principle of subsidiarity was not applicable to the facts of the case. The *SANDU* principle and proviso are aimed at preventing a litigant from circumventing legislation and relying directly on a constitutional right, thereby undermining the constitutionally enshrined role of the legislature. In this case the applicant did not seek to circumvent PAIA or to

¹³¹ Van der Walt 'Normative Pluralism' (note 14 above) at 106–107.

¹³² *My Vote Counts* (note 12 above) at paras 136–149.

¹³³ *Ibid* at paras 155–156.

¹³⁴ *Ibid* at para 162.

¹³⁵ *Ibid* at paras 174–175, footnotes omitted.

undermine the value of constitutional democracy. Instead, the applicant sought to hold the legislature to account for failing to give adequate effect to the right to access to information. In Van der Walt's language, this is the type of case in which the legislature, whether intentionally or by mistake, left a gap in the legislation:

SANDU and *Bato Star* must therefore be applied flexibly; even supposedly codifying legislation will leave gaps and, if the Constitution or the common law provides for such a gap, the next step could be to turn to the Constitution or the common law to fill that gap rather than challenge the legislation. Shifting the problem up to the Constitution or down to the common law before reverting to a constitutional challenge makes sense if a gap in the legislation means that the specific aspect is not covered by the legislative scheme, which means that the subsidiarity principles do not apply and the gap can be filled by application of constitutional provisions or the common law, as illustrated by constitutional review of legislative, executive or judicial acts and by judicial review of administrative action in cases where private bodies exercise public power.¹³⁶

In the circumstances, the minority in *My Vote Counts* correctly held that the principle of subsidiarity was not applicable to the facts of the case because 'the validity of [PAIA] is not at issue'.¹³⁷

In our view, the majority ruling in *My Vote Counts* should be viewed as an extension of the classic *SANDU* proviso based on the doctrine of separation of powers: when a litigant challenges the legislative branch for failing in its constitutional obligation to give effect to a right in the Bill of Rights, subsidiarity and the doctrine of separation of powers demand that a litigant launch a frontal challenge to the validity of the legislation. However, this extension of the *SANDU* proviso could lead to the undesirable situation where the legislature becomes the final arbiter on the boundaries of constitutional rights. The majority's reliance on the doctrine of separation of powers to extend the *SANDU* proviso misconceives the nature of the applicant's case. As mentioned above, instead of impugning the validity of PAIA, the applicant argued that the legislature had failed to fulfil a constitutional obligation. To the extent that this amounts to an infringement of the doctrine of separation of powers, such an infringement is constitutionally sanctioned in s 167(4)(e), which confers exclusive jurisdiction on the Constitutional Court to determine whether the legislature has failed to fulfil a constitutional obligation. As Klare notes, it would be 'paradoxical to entrench judicially enforceable rights in a supreme constitution — a constitution meant to constrain the legislature — and then leave it to the legislature effectively to determine how those rights are to be protected and enforced, subject only to highly deferential judicial review'.¹³⁸

Notwithstanding the differences between the majority and minority regarding the applicability of subsidiarity to the facts of the case, the two judgments constitute an unequivocal confirmation of the constitutional principle of subsidiarity. The *SANDU* principle and proviso, as recognised in *My Vote Counts*, apply equally to other constitutional provisions where the legislature is mandated

¹³⁶ Van der Walt 'Normative Pluralism' (note 14 above) at 109.

¹³⁷ *My Vote Counts* (note 12 above) at para 67.

¹³⁸ Klare (note 28 above) at 143.

to enact legislation to give effect to a constitutional right. This includes the rights to administrative justice. The Constitutional Court has thus, in *My Vote Counts*, arguably foreclosed any debate about how to approach administrative law's threshold questions.

A key difference between *My Vote Counts* and *Motau* is that in the context of access to information cases our courts have not developed a justiciable principle akin to legality – that is, a principle emerging from the constitutional values by which litigants can assert a claim for access of information, and which could usurp the role of s 32 of the Constitution and PAIA as the principle of legality does in relation to s 33 and PAJA. Were such a principle to emerge, it might for instance be couched as ‘the principle of transparency’, another aspect of the rule of law. But the absence of such a principle makes a difference. In *My Vote Counts* it meant that the important role of the legislature in the scheme of separation of powers was explicitly recognised and confronted by the Court. As we discuss above, in *My Vote Counts* both the majority and the minority tested the legislation enacted to give effect to s 32 against s 32. For the majority, the absence of a frontal challenge to PAIA meant that the applicants were unsuccessful. For the minority, a gap in the legislation meant that the applicants had to succeed. In administrative-law cases, when courts flout subsidiarity and ignore the existence of a constitutional continuum of accountability, resort to the principle of legality has the effect of cutting the legislature out of the equation entirely, as PAJA's validity or otherwise is simply overlooked. This difference, for us, reveals that the correct application of subsidiarity is all the more important in administrative-law cases, if the value of democracy is to be upheld.

Further, *My Vote Counts* does what *Motau* failed to do by explicitly offering a coherent and principled method for choosing between the PAJA, s 33 and the principle of legality as potential sources of administrative review. This means that where a litigant seeks to enforce the rights to administrative justice, resort must first be had to the PAJA. Only where the validity of the PAJA (or other original legislation) is challenged may the s 33 rights be invoked directly. Moreover, only in circumstances where the conduct does not amount to administrative action, and provided the PAJA is not found to be inconsistent with s 33 of the Constitution, may resort be had to the principle of legality as a safety net to ensure that the conduct in question does not escape constitutional scrutiny.

IV CONCLUSION: TIME TO STEM THE TIDE

Recent administrative-law jurisprudence of the High Court and Supreme Court of Appeal reveal that, notwithstanding *Motau* and *My Vote Counts*, a lack of clarity persists as to how to address the threshold questions whether and when legality or PAJA should be invoked as the basis to review public power that might amount to administrative action. For instance, in *Minister of Education for the Western Cape v Beauvallon Secondary School*, Leach JA held, in judicial review proceedings concerning a decision to close a number of schools, that it was unnecessary to determine

whether PAJA was applicable.¹³⁹ Dismissing the importance of subsidiarity in one fell swoop, he made the following startling statement:

I am aware that as a rule a court considering the review of a decision of a public official should determine whether or not the proceedings are governed by PAJA. But I do not believe that rule to be rigid and inflexible, as it is indeed now well established that even in cases where PAJA is not of application, the principle of legality may be relied upon to set aside an executive decision made not in accordance with the empowering statute. And in the present case the statutory incorporation into s 33(1) of the Schools Act of a notice and comment procedure essentially the same as that envisaged by s 4(3) of PAJA renders superfluous any attempt to pigeonhole the decision to close the schools as either executive or administrative in nature.¹⁴⁰

A number of High Court judgments have followed suit. One notable example is *Aboobaker NO v Serengeti Rise Body Corporate* where Steyn J, reviewing a decision of the eThekweni Municipality to rezone a property and approve building plans for a development in Durban, failed to answer administrative law's threshold questions, bypassing PAJA and applying legality to the dispute simply because legality had not previously been 'ruled out'.¹⁴¹ In these decisions PAJA is treated as an inconvenience – something the judge need not 'dwell on'.¹⁴² As Cachalia JA observed in 2016 when writing for a majority of the Supreme Court of Appeal in *State Information Technology Agency SOC Ltd v Gijima Holdings (Pty) Ltd*, this is jurisprudentially unacceptable.¹⁴³ In this case the State Information Technology Agency (SITA) applied to have its decision to conclude a contract with Gijima declared unenforceable for lack of compliance with public procurement requirements. In doing so, however, SITA relied on the principle of legality rather than the PAJA, arguing that the latter does not apply when an organ of state is seeking to undo its own decisions. Cachalia JA held:

It is well established that a decision by a state entity to award a contract for services constitutes administrative action in terms of s 1 of PAJA. Once this is accepted, there is no good reason for immunising administrative decisions taken by the state from review under PAJA.¹⁴⁴

However, a minority of the Supreme Court of Appeal found the majority's insistence on applying PAJA to be unduly 'formalistic'.¹⁴⁵ For the minority,

¹³⁹ *Minister of Education, Western Cape and Another v Beauvallon Secondary School and Others* [2014] ZASCA 218, 2015 (2) SA 154 (SCA) at para 16.

¹⁴⁰ *Ibid.*

¹⁴¹ *Aboobaker NO and Others v Serengeti Rise Body Corporate and Another* [2015] ZAKZDHC 54, 2015 (6) SA 200 (KZD) at para 8. See also *Relmar Holdings (Pty) Ltd v Minister of Agriculture, Forestry and Fisheries and Another* [2015] ZAWCHC 103, 2015 JDR 1546 (WCC) at para 73 and *Gidani (Pty) Ltd v Minister of Trade and Industry and Others* [2015] ZAGPPHC 457, 2015 JDR 1471 (GP) ('Gidani').

¹⁴² *Gidani* (note 141 above) at para 58. See also *Malema and Another v Chairman, National Council of Provinces and Another* [2015] ZAWCHC 39, 2015 (4) SA 145 (WCC) at para 47.

¹⁴³ *State Information Technology Agency SOC Ltd v Gijima Holdings (Pty) Ltd* [2016] ZASCA 143, 2017 (2) SA 63 (CC) ('Gijima') at para 16, and see also paras 36–37. This was also acknowledged in *Comair Ltd v Minister of Public Enterprises and Others* [2015] ZAGPPHC 361, 2016 (1) SA 1 (GP) at para 22.

¹⁴⁴ *Gijima* (note 143 above) at para 16.

¹⁴⁵ *Ibid* at para 55.

notwithstanding the administrative nature of the conduct under scrutiny, legality could be invoked if this was what justice demanded.¹⁴⁶

The minority's approach in *Gijima* is problematic.¹⁴⁷ As we have illustrated above, answering administrative law's threshold questions by applying subsidiarity theory affirms the value of democracy and the separation of powers: it gives recognition to the legislature responsible for enacting PAJA. It is eminently more appropriate to test exercises of public power that amount to administrative action with reference to a statute enacted by the legislature (PAJA) than through judge-made law conceived with reference to the flexible principle of legality. At the same time, subsidiarity theory upholds constitutional supremacy. It does so first because it demands the application of a constitutionally mandated statute, PAJA, when that statute is applicable. It places that conduct on a continuum of constitutional accountability through the application of direct norms specifically intended to test public power that amounts to administrative action. Secondly, it allows exercises of power to be tested further along the continuum of accountability through the application of more general constitutional norms when PAJA is not applicable.

Thus, the judicial refusal to answer administrative law's threshold questions is significant for at least two reasons: it amounts to disdain for the legislature's legitimate role in the scheme of separation of powers and it amounts to a rejection of constitutional supremacy. These values have been recognised not only in early Constitutional Court jurisprudence such as *Mhlungu*, *Zantsi* and *New Clicks*, but also more recently in *Motau* and forcefully in *My Vote Counts*. We therefore believe that it is time to stem the tide of the judicial refusal to answer administrative law's threshold questions: subsidiarity theory offers a coherent basis upon which to do so.

¹⁴⁶ Ibid at para 58.

¹⁴⁷ For commentary on the minority's approach see Danie Brand, Melanie Murcott & Werner van der Westhuizen 'Administrative Law' 2016 (3) *Juta's Quarterly Review of South African Law* (forthcoming) at para 2.1.1.

Administrative Action, the Principle of Legality and Deference – The Case of *Minister of Defence and Military Veterans v Motau*

Andrew Konstant*

I INTRODUCTION

*Minister of Defence and Military Veterans v Motau and Others*¹ marks a significant stride in administrative law, the likes of which we have not seen from the Constitutional Court in some time. The judgment tackles in impressive depth two of the fundamental issues in the judicial scrutiny of exercises of public power. The first issue is the complex definition of ‘administrative action’ in the Promotion of Administrative Justice Act 3 of 2000 (PAJA). Secondly, the judgment tackles the difficult question whether procedural fairness exists as a self-standing requirement of the principle of legality. Both of these concerns have surfaced in several previous judgments and have slowly and incrementally evolved into a daunting field of seemingly contradictory judgments. The judgment in *Motau* fleshes out these two questions and puts forward a framework to better guide future judicial decision-making.² However, in doing this, it raises some challenges of its own.

This is not entirely surprising as the judgment covers a great deal of ground. Briefly, the interesting features of the judgment include the use of the principle of legality to achieve a measure of deference with respect to administrative decisions, the content of the principle of legality, and the appointment and dismissal powers of a Minister in charge of an administrative agency. Instead of an assessment of the Court’s entire reasoning, this case note will tackle only one of its more novel features. More specifically, I wish to explore the Court’s use of the principle of legality for the purpose of showing deference towards a decision that it feels would be inappropriate to subject to more searching judicial scrutiny. In doing so,

* Researcher, South African Institute of Advanced Constitutional, Public, Human Rights, and International Law, a Centre of the University of Johannesburg. I am grateful to Cora Hoexter for helpful feedback on a draft of this article. I also benefited from discussions with David Bilchitz, Michael Bishop, Khomotso Moshikaro, Melanie Murcott and other participants at the Constitutional Court Review VII conference, as well as with Nurina Ally. Of course, all remaining errors are my own.

¹ [2014] ZACC 18, 2014 (5) SA 69 (CC), 2014 (8) BCLR 930 (CC) (*‘Motau’*).

² See A Konstant ‘Administrative Action and Procedural Fairness – *Minister of Defence and Military Veterans v Motau*’ (2016) 133 *South African Law Journal* 491 (broader discussion of the impact of the judgment).

however, I am unavoidably drawn into a discussion of the relationship between the principle of legality and PAJA more broadly.

The Court's approach reveals a subtle shift in the way it applies the principle of legality as a pathway of review. Prior to *Motau*, the Court relied on the principle of legality to achieve some measure of accountability in respect of decisions of the legislature and executive that did not constitute 'administrative action' in terms of PAJA. In order to do so, however, the Court began an aggressive campaign to expand the scope of the principle by filling it with the grounds of review ordinarily found in PAJA review. This often gave the impression that the Court viewed the principle of legality as simply a surrogate for PAJA review that was far easier to use. *Motau* is the Court's attempt to correct this impression and make explicit the difference between review under the principle of legality and review under PAJA. The Court adds to a list of distinguishing features of legality and PAJA review by pointing out that the former is the more 'appropriate' method to use in instances where greater deference is required.³

Before delving into the facts of the judgment, it is important to note the shifting nature of the jurisprudential terrain on which the judgment rests. The principle of legality has become a pragmatic tool in the hands of the courts and has, over time, been filled systematically with various grounds of review that are ordinarily found in PAJA. At this point, it is difficult to determine just which grounds of review fall outside the ambit of the principle of legality and remain exclusive components of PAJA review. However, the point to be made here is that the assessment below, and indeed any discussion of the relationship between review under the principle of legality and PAJA review, is subject to the courts' continued manipulation of the principle of legality. As such, any assessment of its content, use or worth will require revision should the Court continue to expand the ambit of the principle of legality.

II FACTS

The case relates to a decision of the Minister of Defence and Military Veterans to remove two directors, the Chairperson and Deputy Chairperson, from the board of the Armaments Corporation of South Africa (SOC) Ltd (Armcor). Armcor is a state-owned entity, and its governing statute is the Armaments Corporation of South Africa Ltd Act 51 of 2003 (the Armcor Act). In terms of the Armcor Act, Armcor falls under the governing jurisdiction of the Minister of Defence. These two directors, General Moreti Motau and Refiloe Mokoena, had their board membership terminated by the Minister after their failure to attend meetings arranged by the Minister for the purpose of dealing with certain procedural issues. This was done in terms of s 8(c) of the Act, which permits such termination on the showing of good cause. The Minister's reasons for terminating the directors' appointments included the lack of expeditious progress with several procurement projects as a result of the board's ineptitude, the failure of the board to conclude a service level agreement with the Department of Defence and Military Veterans as required by s 5 of the Armcor Act and, finally, that the Minister had received

³ *Motau* (note 1 above) at para 43.

several complaints about Armscor from members of the defence industry which she took as evidence of a breakdown in the relationship between the organisation and the industry. In summary, the Minister reasoned that the two directors had not acted in the best interests of the government.⁴

Following the decision of the Minister, both directors took the decision on review on the basis that it was ‘unlawful, unconstitutional and invalid’.⁵ The matter was first heard in the High Court, where Legodi J held that the decision was indeed administrative in nature and therefore fell within the purview of PAJA.⁶ The High Court found that the Minister’s decision was based on an error of law, was procedurally unfair, and that the Minister had acted on the basis of an ulterior motive. The matter was then taken on appeal by the Minister to the Constitutional Court.

This case note will examine the Court’s approach to the definition of ‘administrative action’ as well as the procedural fairness requirements that may flow from the principle of legality.

III ADMINISTRATIVE ACTION

Khampepe J, writing for the majority, admirably sets out several bases on which to decide whether the exercise of public power constitutes ‘administrative action’. Each basis is accompanied by a caution that indicates that no hard-and-fast rule can be extrapolated in order to dispense easily with this assessment. Therefore, as before, the assessment is context-dependent.⁷ The Court begins with the definitional hurdle that must be overcome for the application of PAJA to the decision under scrutiny. Section 1 of PAJA requires that the decision be ‘administrative in nature’. This is, the Court states, a valuable, albeit circular, step in that it forces the reviewing court to closely examine the decision in question.⁸ It also confirms that the primary issue at hand is the nature of the decision and not the identity of the decision-maker.⁹

The Court sets out several indicators that may prove helpful in determining the nature of the power. The first is the source of the power. Where a power flows directly from the Constitution, one could deem the power to be executive in nature. When a power is sourced in legislation, it is likely to be administrative in nature.¹⁰ Secondly, substantial constraints on the power would be an indication that the power is administrative in nature.¹¹ Finally, the court states that the nature of the power can be determined with reference to the appropriateness of subjecting the power to the stricter form of judicial scrutiny represented by the edifice of administrative law contained in PAJA.¹²

⁴ *Motau* (note 1 above) at para 14.

⁵ *Ibid* at para 17.

⁶ *Motau and Another v Minister of Defence and Military Veterans and Another* Case No: 51258/13 (TPD).

⁷ *President of the Republic of South Africa and Others v South African Rugby Football Union and Others* [1999] ZACC 11, 2000 (1) SA 1 (CC), 1999 (10) BCLR 1059 (CC) (‘SARFU’) at para 143.

⁸ *Motau* (note 1 above) at para 34.

⁹ *Ibid* at para 36.

¹⁰ *Ibid* at para 39.

¹¹ *Ibid* at para 41.

¹² *Ibid* at para 43.

This is the major point of departure for Jafta J, who wrote the dissenting judgment. For Jafta J the matter is not nearly as complex and rests largely on the notion that administrative action is involved wherever the implementation of legislation takes place.¹³ As such, unless the governing legislation explicitly or implicitly indicates otherwise, the primary determinative factor is whether the decision constitutes the implementation of legislation.¹⁴ In the present case, the Minister was clearly acting in terms of governing legislation, and therefore her decision constituted administrative action. From here, the prescription is fairly simple: procedural fairness is a requirement in the exercise of administrative powers and thus the Minister was in breach of this requirement when she did not afford the directors an opportunity to be heard before dismissing them.¹⁵

The definition of administrative action and the complexity of the inquiry involved in establishing the nature of public power have created a great deal of difficulty for the courts.¹⁶ Jafta J's approach simplifies the enquiry by arguing that the relevant issue is whether the action constitutes the implementation of a statute. It is arguable, however, whether such a simplistic approach can adequately incorporate all the exercises of public power that ought to be scrutinised in terms of the grounds of review provided for in PAJA, or be enough to justify any distinction between the legality review and PAJA review.

Khampepe J, on the other hand, found that the decision to remove the directors constituted executive action. In reaching this conclusion, Khampepe J turned to the Armscor Act and made three findings. First, the Minister's power in terms of s 8(c) is an adjunct to her power to formulate defence policy. The nature of the policy-making power that the Minister wields is wide and rather abstract, and does not involve the details of the day-to-day operations of the corporation. These details are instead left to the board. As a result, her involvement is limited to guiding the strategic direction of the organisation by appointing and dismissing leaders of the corporation.¹⁷ In conclusion, the majority judgment states that even though the appointment or dismissal of management is not itself policy creation, it is the means through which the Minister 'gives direction in the vital area of military procurement, and is therefore an adjunct to her executive policy formulation function'.¹⁸

The second point raised by Khampepe J is that the power is not a low-level bureaucratic power. In other words, it is not a power that simply involves the application of policy in the daily functioning of the state. The power operates exclusively between the Minister and high-level managers and allows for the supervision of these managers. This therefore suggests the exercise of executive power rather than administrative power.¹⁹ The third and final point that Khampepe J relies on is that the power is fairly unconstrained. The Minister need only rely on good cause to dismiss a board member. Thus, the Minister

¹³ Ibid at para 106.

¹⁴ This is a fundamental factor laid down in *SARFU* (note 7 above) at para 142.

¹⁵ *Motau* (note 1 above) at para 127.

¹⁶ C Hoexter *Administrative Law in South Africa* (2nd Edition, 2012) at 195.

¹⁷ *Motau* (note 1 above) at para 48.

¹⁸ Ibid at para 48.

¹⁹ Ibid at para 49.

exercises broad discretion with respect to appointment and dismissal, which again is indicative of executive rather than administrative power.²⁰

A The Relationship Between the Principle of Legality and PAJA

The *Motau* judgment is the Court's re-entry after a long hiatus into the complicated jurisprudence dealing with the lawful exercise of public power. Since the courts' adoption of the principle of legality, there has been the need to organise the two main pathways of judicial review on a more principled foundation. As in the case of s 33 of the Constitution and PAJA, the principle of subsidiarity could be used to achieve this.²¹ The Court in *Motau* certainly makes an effort towards organisation; however, instead of exclusively using the principle of subsidiarity, the Court suggests that one organising principle may be the need to show deference to a decision of the executive.²² Specifically, Khampepe J suggests that defining the decision as executive in nature can be justified by the need to 'show the Executive a greater level of deference'.²³ Khampepe J goes on to state that

the Court has found that administrative law review is not appropriate where the power under consideration: is legislative in nature and influenced by political considerations for which public officials are accountable to the electorate; is based on considerations of comity or reciprocity between South Africa and foreign states, involving policy considerations regarding foreign affairs; is closely related to the special relationship

²⁰ Ibid at para 50.

²¹ See C Hoexter 'A Rainbow of One Colour? Judicial Review on Substantive Grounds in South African Law' in H Wilberg & M Elliott (eds) *The Scope and Intensity of Substantive Review: Traversing Taggart's Rainbow* (2015) 183 (author makes such a call). See also *Albutt v Centre for the Study of Violence and Reconciliation and Others* [2010] ZACC 4, 2010 (3) SA 293 (CC), 2010 (5) BCLR 391 (CC) at para 82 ('*Albutt*') (provides an interesting counterpoint to this when the Court states that 'judicial policy' may prevent a court from even asking whether PAJA should be applied where the principle of legality is capable of resolving the matter). The Court in *Motau* does not deal with this assertion and has perhaps retreated from this position. One could, however, offer an argument against the designation of the principle of legality operating at a higher level of generality than s 33 of the Constitution. For instance, Fowkes argues that the principle of legality operates as a rule distinct from the principle found in s 1(c) of the Constitution. See J Fowkes 'Founding Provisions' in S Woolman & M Bishop (eds) *Constitutional Law of South Africa* (2nd Edition, OS, 2008) ch 13. If true, establishing that the principle of legality operates at a higher level of generality in relation to s 33 would be contestable. Likewise, it could be argued that the principle of legality acts as a *de facto* right to lawful public action. Such an argument would obviously stand in tension with the Court's explicit statement that the Constitution's founding provisions are not rights. *Minister of Home Affairs v National Institute for Crime Prevention and the Reintegration of Offenders and Others* [2004] ZACC 10, 2005 (3) SA 280 (CC), 2004 (5) BCLR 445 (CC) ('*NICRO*') at para 23. However, as Fowkes attempts to show, it is not easy to argue that the Court has been consistent in abiding by its statement in *NICRO*. Should such an argument prove persuasive, it would be difficult to determine whether such a right was more or less general than s 33. Ultimately, making an argument that subsidiarity can adequately organise the various pathways of review will require a convincing explanation of the provenance of the principle of legality and a general taxonomy of review.

²² I treat the two issues as largely distinct from one another. In my view, the intensity of a standard of review has little to do with the generality with which it is articulated. The principle of subsidiarity, as far as we know, does nothing more than tell us which norm should be applied where more than one is applicable. It does not necessarily designate one as more intensive than the other. Thus this note discusses the distinction in intensity of review without dealing with the Court's use of the principle of subsidiarity.

²³ *Motau* (note 1 above) at para 43.

between the president and the Director-General of a security agency; or involves the balancing of complex factors and sensitive subject matter relating judicial independence.²⁴

This is not the first time a court has made the argument that the use of the principle of legality is justified on the basis that it is more deferential than PAJA review. In 2014, in *NDPP v FULL*, the Supreme Court of Appeal (SCA) attempted to end the debate whether a decision not to prosecute is reviewable under PAJA.²⁵ The debate emerges from uncertainty whether PAJA, when excluding ‘decisions to prosecute’, also excludes decisions *not* to prosecute from the concept of administrative action. The central question is whether there is any reason to treat the two forms of decision differently. Brand JA held that the two policy justifications that exist for denying or limiting the courts’ review powers with respect to a decision to prosecute apply equally to decisions not to prosecute. The first is safeguarding the ‘independence of the prosecuting authority by limiting the extent to which review of its decisions can be sought’.²⁶ The second justification is the width of the discretion exercised by the prosecuting authority and the ‘polycentric character that generally accompanies its decision-making’.²⁷ The implicit argument here can only be that the principle of legality is a form of review that poses less risk to the independence of the prosecuting authority than PAJA review does.

With *Motau* following suit, the Court has introduced the notion that the nature of a power can be determined, albeit in part, with reference to the appropriateness of subjecting that power to the stricter form of judicial scrutiny in terms of PAJA. The implicit argument here is that there is a category of decisions to which the courts should apply a light touch or deferential form of review. In other words, we have the Court hinting that PAJA review does not possess within it the necessary capacity for deference with respect to decisions of the executive branch of government. The Court, in adopting this line of reasoning, suggests an instrumentalist use of the principle of legality in order to achieve its preferred degree of deference. There are roots of this sentiment in previous cases. For instance, in *New Clicks*, Sachs J states that ‘judicial review of subordinate legislation can be more effectively and robustly done if not forced to tip-toe on the narrow pedestal appropriate for reviewing administrative acts’.²⁸ On the other hand, examples of judgments exist in which the only distinction pointed out is that the principle of legality applies to a wider set of public powers, as opposed to PAJA review, which is restricted to decisions defined as ‘administrative action’.²⁹

Scholars have also argued that the principle of legality is better suited to more deferential review and thus more appropriate for the review of executive decisions.³⁰ It is, however, important to note that these suggestions do not include

²⁴ Ibid.

²⁵ *National Director of Public Prosecutions and Others v Freedom Under Law* [2014] ZASCA 58, 2014 (4) SA 298 (SCA).

²⁶ Ibid at para 25.

²⁷ Ibid.

²⁸ *Minister of Health and Another NO v New Clicks South Africa (Pty) Ltd and Others (Treatment Action Campaign as Amici Curiae)* [2005] ZACC 14, 2006 (2) SA 311 (CC), 2006 (1) BCLR 1 (CC) at para 609.

²⁹ *Albutt* (note 21 above) at para 49.

³⁰ Hoexter *Administrative Law* (note 16 above) 124.

the assertion that the distinction in depth of review should be used by the courts as reasons for classifying a decision as administrative in nature or not. And most importantly, as in Hoexter's case especially, the suggestion is caveated by the requirement that the principle of legality contains only the grounds of review that reflect the bare minimum of what is required of administrative decisions.³¹ Therefore, the more grounds of review that the principle of legality adopts, the less able it is to play the role of a more deferential pathway of review.

The result of the Court's adoption of the distinction between the intensity of review of the principle of legality and PAJA review is that such a distinction can be and has been used as a partial justification for the organisation of the two forms of review. The other potential justification is the principle of subsidiarity. *Motau's* organisational principle can thus be briefly described as the following: first the court considers whether the decision is administrative in nature. In doing so it may consider, among other factors, whether it wants to show the decision a degree of deference. If, after considering these factors, the court finds that the decision is not administrative in nature, it may well have to decide whether the decision is executive or legislative and will apply the principle of legality. This sequence of reasoning reveals several important changes in the role that the principle of legality plays in judicial review.

IV THE PRINCIPLE OF LEGALITY

The *Motau* judgment reflects the Court sharpening its understanding of the relationship between the principle of legality and PAJA. In order to make this clear, it may be helpful to lay out the short history of the principle and its early use by the Court.³² Any such exposition must begin with s 1(c) of the Constitution in which the principle is anchored. This founding provision ensures that, as a constitutional democracy, the state holds the Constitution as the supreme law and upholds the rule of law. It is the concept of the 'rule of law' specifically from which the Court initially drew in constructing the principle of legality in its doctrinal form.³³

This history of the principle has been well documented and there is a reasonably consistent narrative. The origin of the principle was explained in *Fedsure*.³⁴ The

³¹ C Hoexter 'The Future of Judicial Review in South African Administrative Law' (2000) 117 *South African Law Journal* 484, 507.

³² For a fuller depiction, see C Hoexter 'The Principle of Legality in South African Administrative Law' (2004) 4 *Macquarie Law Journal* 165, 174.

³³ See F Michelman 'The Rule of Law, Legality and the Supremacy of the Constitution' in S Woolman & M Bishop (eds) *Constitutional Law of South Africa* (2nd Edition, OS, 2008) ch 11 at 15 (Description of the novelty of instrumentalising the rule of law or the principle of legality in this manner).

³⁴ *Fedsure Life Assurance Ltd and Others v Greater Johannesburg Transitional Metropolitan Council and Others* [1998] ZACC 17, 1999 (1) SA 374 (CC), 1998 (12) BCLR 1458 (CC) ('*Fedsure*') at para 59. Cf Fowkes (note 21 above) (Argues that the principle of legality has a more complex past than is ordinarily thought. Fowkes' reading of *Fedsure* is that the Court's reasoning was based on the Interim Constitution which did not have a provision like s 1(c). Thus, in Fowkes' view, the Court in that case relied on the general notion of 'constitutionalism' to anchor the principle of legality in the Interim Constitution. Later cases though have had the benefit of a 'rule of law' provision and have therefore moved away from the construction in *Fedsure*.)

Court found that the principle was ‘implicit in the Constitution’ and reflected the notion that ‘the exercise of public power is only legitimate where lawful’.³⁵ The principle gained further content in *Pharmaceutical Manufacturers* where the Court held that rationality was a ‘minimum threshold requirement applicable to the exercise of all public power’.³⁶ More recently the Court and the SCA have added aspects of procedural fairness and the obligation to give reasons.³⁷

The mooring of the principle of legality in the Constitution does not explain why the Court turned to the principle of legality as an alternative avenue of review to PAJA. The simple answer largely lies in the Courts’ aversion to the definition of ‘administrative action’ in PAJA on one hand, and the inability to allow a class of decision to be unreviewable on the other. The rationale of placing the definition in PAJA was to narrow the applicability of the legislation. It was evident that the Final Constitution provided many other means to deal adequately with ‘non-administrative’ decisions and that PAJA need only concern itself with administrative decisions of a certain type.³⁸ The definition, however, proved to be too complex for the courts to navigate. As a result, it became necessary for the Court to develop another avenue of review equipped with the grounds necessary to hold non-administrative exercises of public power to account. A sub-textual reason was certainly to permit courts to avoid the burdensome reasoning that lurks behind the definition of administrative action in PAJA.

The reason the Court resorted to the principle of legality has been described as allowing for a pattern of avoidance by the courts with respect to the definition of ‘administrative action’ in PAJA.³⁹ Some judges have effectively presented review in terms of the principle of legality and PAJA review as substitutable. They have argued that there is no inflexible rule dictating that courts confront the definition of ‘administration action’.⁴⁰ Instead, courts may simply assume the application of the principle of legality without worrying much about determining the nature of the decision.⁴¹ Other forms of avoidance include the direct application of s 33 in place of PAJA. This strategy has come under criticism from scholars whose concern is respect for the principle of subsidiarity.⁴²

³⁵ *Fedsure* ibid at paras 56 and 59.

³⁶ *Pharmaceutical Manufacturers Association of South Africa and Another; In re Ex Parte President of the Republic of South Africa and Others* [2000] ZACC 1; 2000 (2) SA 674 (CC), 2000 (3) BCLR 241 (CC) at para 90.

³⁷ *Albutt* (note 21 above); *Democratic Alliance v President of South Africa and Others* [2012] ZACC 24, 2013 (1) SA 248 (CC), 2012 (12) BCLR 1297 (CC); and *Judicial Service Commission and Another v Cape Bar Council and Another* [2012] ZASCA 115, 2013 (1) SA 170 (SCA).

³⁸ See C Hoexter ‘The Constitutionalism and Codification of Judicial Review in South Africa’ in C Forsyth et al (eds) *Effective Judicial Review: A Cornerstone of Good Governance* (2010). See also J Klaaren & G Penfold ‘Just Administrative Action’ in S Woolman & M Bishop (eds) *Constitutional Law of South Africa* (2nd Edition, OS, 2008) Chapter 63 at 63–65.

³⁹ Hoexter (note 16 above) at 131.

⁴⁰ See, eg, *Minister of Education for the Western Cape and Another v Beauvallon Secondary School and Others* [2014] ZASCA 218, 2015 (2) SA 154 (SCA), [2015] 1 All SA 542 (SCA) at para 16; *Democratic Alliance* (note 37 above) at para 12.

⁴¹ The same form of reasoning was employed in *Minister of Home Affairs and Others v Scalabrini Centre, Cape Town and Others* [2013] ZASCA 134, 2013 (6) SA 421 (SCA), [2013] 4 All SA 571 (SCA) (‘*Scalabrini*’).

⁴² See, eg, C Hoexter ‘“Administrative Action” in the Courts’ (2006) *Acta Juridica* 303.

It has become clear that the Court has an instrumental aim in the use of the principle of legality. In *SARFU*, the Court states that just because a decision is classified as executive in nature, this does not mean that there are no constraints placed upon it.⁴³ Added to this is the Court's expressed position that the principle of legality is not a static doctrine and will evolve with changing circumstances. This statement was made in the course of expanding the scope of the doctrine, and it would not be unreasonable then to assume that the form of evolution that the Court had in mind was the principle's expansion. This seems to suggest that the Court's desire to fill the principle of legality with the tools of PAJA review is being driven by a need to constrain executive decisions.⁴⁴ In other words, the stronger a court's desire to restrain executive decision-making, the more expansive the principle of legality will become.⁴⁵

The purpose of this section is to show that the courts have changed the way in which the principle of legality is being viewed and used. It has been considered a catch-all that had, as its purpose, holding the executive or legislature to account.⁴⁶ Coupled with the expansion undertaken by the Court, one might be tempted to think that the principle of legality was as sharp a tool for holding other arms of government to account as PAJA review. In *Motau*, the Court has sought to alter this perception. It has made the most explicit movement towards viewing the principle of legality as an instrument to show deference to executive or legislative decisions.⁴⁷ When once it appeared that the Court would rely on almost any justification sourced from the principle of legality or 'rule of law' to hold the executive accountable, it is now positioning the principle of legality as a weaker and less invasive form of review.⁴⁸

The Court indicates quite clearly in *Motau* that one can expect a less intrusive review under the principle of legality than under PAJA review. We could ask whether the jurisprudential 'supporting structures' exist within our law for such an assumption, whether good or bad. In order to attempt an answer, we must identify the features of the two pathways of review that lead to a difference in the intensity of the review when either is applied. The first possible candidate is that there exists a difference in the width of deference available in each. Perhaps, when using the principle of legality, a judge feels more constrained in assessing the decision than if she applied PAJA. The other candidate is the set of grounds of review available in each pathway of review. If we can determine that they differ

⁴³ *SARFU* (note 7 above) at para 148.

⁴⁴ A Price 'The Evolution of the Rule of Law' (2013) 130 *South African Law Journal* 657.

⁴⁵ See Price (ibid) for a useful account of the expansion of the grounds of review under the principle of legality and their counterparts in administrative law review.

⁴⁶ See Michelman (note 33 above) at 16–23 (argues that a significant purpose of the principle of legality in the early stages of the Court's jurisprudence was as a method to secure its jurisdiction over the Supreme Court of Appeal).

⁴⁷ This is also not surprising given the Court's movements towards affording the public institutions a wider berth when reviewing their actions. This is typified in the majority's position in the recent decision in *My Vote Counts NPC v Speaker of the National Assembly and Others* [2015] ZACC 31, 2016 (1) SA 132 (CC) ('*My Vote Counts*') and, in particular, the judgment's stance on deference.

⁴⁸ More recently the SCA has also opted to recast the principle of legality as a milder form of review by describing it as a 'measure of last resort'. See *State Information Technology Agency Soc Ltd v Gijima Holdings (Pty) Ltd* [2016] ZASCA 143, [2016] 4 All SA 842, 2017 (2) SA 63 (SCA) at para 38.

in either identity or use, then perhaps there is justification for believing that one form of review is more deferential than the other.

A Value of *Motau's* Organisational Method

Now that we have dealt with the implications of the judgment with respect to the principle of legality, the question is whether the distinction between the two paths of review on the basis of their respective standards is sensible. The judgment, understandably given that the matter did not turn on this point, did not elaborate on the specific aspects of the two forms of review that yield different levels of intensity. This section will proceed to suggest the possible mechanisms within each form of review that may support the Court's conclusion and the weaknesses that each possesses. Before this is done, however, there are some initial problems with the Court's attempt to deal with the relationship between the principle of legality and administrative law review in this manner.

The judgment implies that the courts should define a decision as executive on the basis that there may be a desire to apply some degree of deference to the decision. Should this motivation exist as a self-standing reason for characterising a decision as executive in nature, it would be a rather unprincipled way of defining executive action. Further, it conflates the nature of the power with the respect that a court should show it. These are two separate steps with different reasons that guide the determination of each. The nature of the power refers to whether the power falls into one of the four defined categories: judicial, executive, legislative or administrative. It involves a determination with reference to the features of each to determine whether the decision under review meets a particular definitional threshold. The level of deference is a separate concern. Deference with respect to each of these categories exists on a sliding scale. Reasons for deference may operate differently within each category, and it would be difficult to ascertain a standard level of deference in every case of executive, judicial, legislative or administrative decisions. Degrading the distinction between these two steps would make an already challenging process of determining the nature of a decision and the requisite judicial approach all the more difficult.

Beyond this initial issue, there is the question of whether such a distinction between legality and PAJA review can be sustained. In order to determine whether this is possible, we must be able to draw a line between the level of intrusion by a court when using the principle of legality as opposed to PAJA review. This distinction should have some positive basis. We should be able to justify it by pointing to inherent features of the pathways that make one always more appropriate in certain instances than the other. Before trying to locate these features, it will be instructive to look at the features that we already know exist. The Court, in making this distinction, gestures towards deference. This is not a new concept, even if it remains somewhat undeveloped in our administrative law jurisprudence, and it may pose problems for the Court's perspective on the levels of intensity in both forms of review.

B Deference in General

The two pathways being distinguishable on the basis of the appropriateness of their application to decisions of the executive implies that there exists a distinction in the level of deference available in each pathway. The notion or doctrine of deference is not very well developed in South African jurisprudence despite the urging of some.⁴⁹ The consequence of *Motau* is to make apparent the deficiencies of an un-nuanced notion of deference and the concomitant lack of attention paid to the standards of review attached to each form of review. For the Court to have assumed that legality and PAJA review differ in terms of their respective levels of intrusion, there must presumably have been some sophisticated and unexplained understanding of deference that justifies the Court's reasoning. The reason for this assumption is that, as far as we know, the deference mechanism that we currently have in administrative law is sufficient to calibrate a court's approach to a decision on the same basis that justifies a court's review of a decision of the executive. If deference under PAJA review is capable of achieving the same aim the Court has with respect to the principle of legality, it would be difficult to argue that a distinction between the two forms of review exists. To better understand this argument, it would be useful to capture our best understanding of the Court's approach to deference.

Substantial scholarship has been devoted to the need for a self-standing doctrine of deference,⁵⁰ the content of such a doctrine or theory, and its uses in different forms of judicial review.⁵¹ It is not possible in the space available here to sift through this literature or to adequately formulate a doctrine or theory of deference. In place of such an analysis, it may only be necessary to lay out in general what is meant by deference in the judicial setting and the fundamental principles that guide or justify it. These principles will facilitate some understanding of what the Court was trying to achieve in *Motau* and whether these efforts could be successful.

⁴⁹ Hoexter (note 31 above). Cf DM Davis 'To Defer and When? Administrative Law and Constitutional Democracy' (2006) *Acta Juridica* 23. See also PJH Maree & G Quinot 'A Decade and Half of Deference' (2016) *Tydskrif vir die Suid-Afrikaanse Reg* 268 (a recent description of the Court's construction of deference and its discussion in South African academic circles).

⁵⁰ Allan is perhaps the most vocal opponent of a doctrine of deference. Despite this, he is committed to many of the underlying principles of deference and essentially argues that these principles can find better expression by judges reacting to particular circumstances of each case. In other words, there is no need for the establishment of a doctrine that seeks to guide the decision-making of judges in every relevant case. See TRS Allan 'Common Law Reason and the Limits of Judicial Deference' in D Dyzenhaus (ed) *The Unity of Public Law* (2004) 289. See also H Corder 'Without Deference, With Respect: A Response to Justice O'Regan' (2004) 121 *South African Law Journal* 438, 441. Cf A Kavanagh 'Defending Deference in Public Law and Constitutional Theory: A Reply to TRS Allan' (2010) 126 *Law Quarterly Review* 236. It is not necessary for the purposes of the arguments in this note to adopt a position with respect to whether a doctrine of deference is needed or whether we can make do with identifying and relying on the reasons for deference in each particular case.

⁵¹ P Daly *A Theory of Deference in Administrative Law* (2012); M Lewans *Administrative Law and Judicial Deference* (2016); D Dyzenhaus 'The Politics of Deference: Judicial Review and Democracy' in M Taggart (ed) *The Province of Administrative Law* (1997) 279; M Hunt 'Sovereignty's Blight: Why Contemporary Public Law Needs the Concept of "Due Deference"' in N Bamforth & P Keyland (eds) *Public Law in a Multi-Layered Constitution* (2003) 337; J King 'Institutional Approaches to Judicial Restraint' (2008) 28 *Oxford Journal of Legal Studies* 409.

As far as consensus is concerned, most scholars agree that there are several drivers of deference in judicial decision-making. On the one hand we have the constitutional mandate with which each arm of government has been endowed. This may be viewed as some form of plenary power in decision-making of a particular kind or over particular substantive areas. The belief is that, at the very least, each arm of government when acting as a check over the other should respect the plenary power. Included in this principle is the injunction against any arm of government assuming the plenary power of another. This prohibition is usually presented under the heading of the separation of powers. Other, more pragmatic, principles that one can use as justifications for deference focus on the institutional capacity of the courts to decide matters ordinarily governed by other branches of government.⁵² Acting as distinct reasons for deference, they could lead to a scenario in which a court may possess the constitutional mandate to check the power of another arm of government. This mandate, however, does not preclude the court from deferring to that arm of government on the basis of a disparity in the institutional competencies of the two institutions.⁵³ As a result, the justification for deference may rest neatly on an interpretation of the separation of powers, as well as the institutional features of the court and its capacity to decide certain matters.

These institutional reasons for deference have more recently been packaged by Aileen Kavanagh as a 'doctrine' of deference that is guided by the institutional capacity of the courts in adjudicating matters better suited to the deliberation of administrative or executive bodies.⁵⁴ These institutional reasons for deference are broken down into the greater institutional competence, expertise, or constitutional/democratic legitimacy of an administrative body to decide a matter. Kavanagh suggests that, instead of creating a distinction between policy and non-policy laden decisions (the former requiring deference and the latter not), the courts should instead make a distinction between 'the type of policy decision appropriate to the institutional features, competence, and legitimacy of the courts and the type of policy decision that is beyond that competence'.⁵⁵ This interpretation, and the focus on the competency of the court in relation to the subject matter in question, would mean that, for instance, the question of the fishing quotas dealt with in *Foodcorp*⁵⁶ and *Bato Star*⁵⁷ warranted deference of the court not by virtue of the fact that the decision involved an assessment of public

⁵² See J Jowell 'Due Deference under the Human Rights Act' in J Jowell & J Cooper (eds) *Justice/UCL Seminars* (2003). (Jowell argues that these two sets of justifications are distinct and should not be grouped together. Jowell separates the notion of constitutional competence or permissibility from institutional competence.)

⁵³ J Jowell 'Judicial Deference and Human Rights: A Question of Competence' in P Craig & R Rawlings (eds) *Law and Administration in Europe Essays in Honour of Carol Harlow* (2003) 67.

⁵⁴ See also C Sunstein and A Vermuele 'Interpretation and Institutions' (2003) 101 *Michigan Law Review* 885.

⁵⁵ A Kavanagh 'Deference or Defiance? The Limits of the Judicial Role in Constitutional Adjudication' in G Huscroft (ed) *Expounding the Constitution: Essays in Constitutional Theory* (2008) 194, 197.

⁵⁶ *Foodcorp (Pty) Ltd v Deputy Director General: Department of Environmental Affairs and Tourism: Branch Marine and Coastal Management* 2004 (5) SA 91 (C).

⁵⁷ *Bato Star Fishing (Pty) Ltd v Minister of Environmental Affairs and Tourism and Others* [2004] ZACC 15, 2004 (4) SA 490 (CC), 2004 (7) BCLR 687 (CC).

policy (a separation of powers issue), but rather because the decision involved public policy *and* was beyond the ability of the court to make.

Trying to determine whether at least some of the above theory has manifested in case-law is not difficult. In *Bato Star Fishing (Pty) Ltd v Minister of Environmental Affairs and Tourism and Others*,⁵⁸ O'Regan J, writing for the majority, held that deference in the review of administrative power springs from 'the fundamental principle of separation of powers'.⁵⁹ What emerges from this principle is a two-pronged test for determining the degree of deference appropriate in any given circumstance. The first factor is the appropriate role of the executive and the legislature in terms of the Constitution. This factor requires the Court's recognition of the democratic legitimacy of each branch of government and is therefore a reference point for the Court's boundary of intrusion. The second factor that the Court will consider is the institutional competence of the executive or legislature in comparison to its own.⁶⁰ In other words, the Court will remain alert and respond to its own institutional limitation and will not seek to venture too far into terrain for which it does not possess the necessary expertise or knowledge. Within this second element would exist concerns of polycentricity or other reasons typically used to justify the Court's lack of ability to scrutinise the merits of a decision taken by the executive or legislature. In reality, the standard adopted in judicial review is variable depending on the circumstances of each case.⁶¹

If one were to juxtapose PAJA review with its attendant mechanism of deference, alongside a principle of legality which the courts have not explicitly imbued with a notion of deference, one could argue that the principle of legality appears to be the more invasive form of review. The appearance becomes even more stark in the face of the continued expansion of the principle of legality's scope of review.⁶²

The implication of the judgment in *Motau* is that we can now discern two points in the public law review process where the choice for deference may be made. The first is situated in the act of deciding whether the exercise of a specific public power constitutes administrative action and manifests itself by characterising the power as executive in nature. Thus, acting as both reason for and consequence of the choice is the application of the less invasive legality review to the power in question.⁶³ The second location of deference occurs, in theory, within a PAJA review. For the time being, all we know for certain is that a court, after having decided that the decision should be tested against the grounds of review contained in PAJA, may calibrate the intensity of its scrutiny through the device of deference.⁶⁴

⁵⁸ Ibid.

⁵⁹ Ibid at para 46.

⁶⁰ Ibid at para 48.

⁶¹ Hoexter (note 31 above) at 502.

⁶² See L Kohn 'The Burgeoning Constitutional Requirement of Rationality and the Separation of Powers: Has Rationality Review Gone Too Far?' (2013) 130 *South African Law Journal* 810.

⁶³ L Kohn 'Our Curious Administrative Law Love Triangle: The Complex Interplay Between PAJA, the Constitution and the Common Law' (2013) 28 *South African Public Law* 22, 33.

⁶⁴ Hoexter (note 32 above) at 184.

The difficulty that arises as a result of *Motau* is this: If deference is going to act as a reason for opting for a principle of legality review over a PAJA review, and PAJA review itself contains a mechanism for deference, then the deference offered by the principle of legality must be greater than that offered by PAJA review. If this is not the case, and we can achieve just as much deference under PAJA review as we can under the principle of legality, then there is little reason to opt for the principle of legality review. Our problem is that it is not entirely clear how the standards of review in the two forms of review differ. As demonstrated above, the current notion of deference is broad enough that a court can be just as deferential under one form of review as it can be under the other.⁶⁵ As such, one option may be to build in different principles of deference into each form of review where one is a greater constraint on judicial discretion than the other. The alternative option is to make a distinction between the intrusiveness of review on the basis of the grounds of review contained in each pathway. The feasibility of this latter option may take greater explication.

C Grounds of review

Putting aside deference as an option for distinguishing between PAJA review and the principle of legality, perhaps the next methodological option available is to assess how well the Court's distinction can rest on the argument that the two pathways of review each contain a distinct, albeit increasingly similar, set of grounds of review. It is possible that the Court views these distinctions as resulting in the difference in the intensity of review offered by each pathway. The validity of this perception hinges on grounds of review playing the determinative role in calibrating the standard of review attaching to each pathway of review.

At the outset we encounter the first problem. There are substantial overlaps between the grounds of review available under the principle of legality and PAJA review. Coupled with this, the Court has arguably closed the door to the option of distinguishing between the two pathways of review on the basis of differing levels of scrutiny where there are overlapping grounds of review. In *Democratic Alliance v President of the Republic of South Africa and Others*, the Court held that there was no need to believe that the test for rationality in terms of the principle of legality and PAJA review should be any different.⁶⁶ The Court went so far as to say that '[i]t cannot be suggested that a decision that would be irrational in an administrative law setting might mutate into a rational decision if the decision being evaluated was an executive one'.⁶⁷ In other words, the Court stated that the overlapping grounds of review in the principle of legality and PAJA review work in the same way and should lead to the same outcome. This would make it difficult to argue that under a bifurcated model of deference, as suggested above, a court can arrive at two distinct outcomes when the same ground of review is used. This is despite reasons for deference nudging the court in two different directions.

⁶⁵ See C Hoexter 'The Enforcement of an Official Promise: Form, Substance and the Constitutional Court' (2015) 132 *South African Law Journal* 222.

⁶⁶ *Democratic Alliance* (note 37 above) at para 44.

⁶⁷ *Ibid.*

This would then require us to read *Motau* as arguing that the non-overlapping grounds of review did the work in distinguishing the intensity of review in the principle of legality and PAJA review. There is some authority for this reading of the judgment. In *New National Party v Yacoob* J argued that rationality was the more appropriate ground of review to use than reasonableness when testing legislative schemes.⁶⁸ Of course, the distinction between the forms of rationality review used for assessing legislation and public power may dilute the usefulness of *Yacoob* J's holding.⁶⁹ Nevertheless, in his note Price defends the Court's distinction of the two grounds of review on the basis that rationality acts as the 'baseline' standard which all public power (and legislative acts) must meet. Reasonableness on the other hand requires far more of the state as justification for its actions.⁷⁰ Therefore, if aspects of the ground of reasonableness only were available under PAJA review, then that would make this pathway of review more intrusive than the principle of legality.

This, in my view, is a difficult argument to make persuasively. Being most generous, one could argue that the grounds of review play an important role in modulating the intensity of the court's review. This has certainly been argued before. A compelling example of this argument is that the ground of rationality steers a court away from the merits of a decision and thus ensures a more deferential form of review.⁷¹ The same could be said of the ground of procedural fairness which has the explicit aim of reviewing only how a decision was made and not why it was made.⁷² However, even the most ardent believers in the usefulness of the grounds of review as a way to calibrate the intensity of review accept that each ground of review can be applied with a variable standard of review which sharply diminishes their independent ability to dictate the intensity of review.⁷³ The Court has certainly made clear that it does not regard grounds of review as possessing a single standard of intensity and that they can be applied in a manner that affords the decision-maker a variable degree of deference.⁷⁴ This has created an uneven standard of scrutiny applied by the Court in cases which have relied on the principle of legality's ground of rationality.⁷⁵

Aside from the reality that rationality does not offer a fixed standard of review, relying on a merits/procedure distinction as a way of tagging a ground of review as deferential or not, is problematic. On the spectrum of which grounds point

⁶⁸ *New National Party v Government of the Republic of South Africa and Others* [1999] ZACC 5; 1999 (3) SA 191 (CC); 1999 (5) BCLR 489 (CC). The same point is made in *Pharmaceutical Manufacturers* (note 36 above) at para 78.

⁶⁹ See A Price 'The Content and Justification of Rationality Review' in S Woolman & D Bilchitz (eds) *Is This Seat Taken? Conversations at the Bar, the Bench and the Academy about the South African Constitution* (2012) 37, 45.

⁷⁰ As suggested above, this is a common assumption. It seems intuitive that because rationality is an easier standard for the state to meet, it is less intrusive than a ground such as reasonableness.

⁷¹ M Fordham 'Surveying the Grounds: Key Themes in Judicial Intervention' in P Leyland & T Woods (eds) *Administrative Law Facing the Future: Old Constraints and New Horizons* (1997) 184, 189.

⁷² *Ibid.*

⁷³ P Craig 'The Common Law, Shared Power and Judicial Review' (2004) 24 *Oxford Journal of Legal Studies* 237.

⁷⁴ Price (note 69 above) at 54.

⁷⁵ A Price 'Rationality Review of Legislative and Executive Decisions: *Poverty Alleviation Network and Albutt*' (2010) 127 *South African Law Journal* 580.

towards the merits of the decision and which point towards process (the implication being that merit based review is inherently more intrusive than process-based review), lawfulness sits squarely on the merit side.⁷⁶ Given that the core of the principle of legality must be the ground of lawfulness, this should indicate that it is *more* invasive than grounds that, at least prior to the expansion of the principle of legality, are typically only found in PAJA review. This is clearly displayed in English law where courts arrogate to themselves almost unfettered power to decide questions of law.⁷⁷ On the merit/substance formulation, the question of deference seems to swing away from the core conception of 'legality'.⁷⁸ Yet, as has been made clear, lawfulness sits alongside, or even below, rationality as the lowest and easiest bar that public power must meet. It becomes less clear, when we delve into the minutiae, whether courts are more or less deferential when it comes to the ground of lawfulness.

Perhaps an even more fine-grained view of the problem may help to illustrate the difficulty in identifying which ground of review in its application is more or less deferential than another. As an example, take the grounds of lawfulness and reasonableness. With respect to a ground of lawfulness, there is less scope for a court to accept the state's argument that an action is lawful. The court will determine that on its own. In other words, there are fewer reasons to accept or add weight to justifications offered by the state. Yet asking whether an action is lawful is a far narrower question to ask than if the action was reasonable. On the other hand, when asking whether an action was reasonable, the court has the latitude to accept the state's assertion that the action is, indeed, reasonable. In fact, there may be good reasons for the court to do so. The court is thought not to be in a position to gauge for itself whether the action is reasonable. On the other hand, the question of whether an action is reasonable is far broader and more challenging for the state to provide justifications.

So, how then do we clearly, reliably and – most importantly – objectively, decide which threshold of review is actually higher? Take it as probabilities. The state is less probable to meet the higher threshold in reasonableness review, but more probable to fall under the reasons for deference. In lawfulness review, the state is more probable to meet the threshold, but less probable to be saved under reasons for deference. Given an equal number of similar cases going through each avenue of review, is it possible, *ex ante*, to tell which threshold will yield the most wins for the state?

What becomes clear from this view of the grounds of review is that the role of deference is far greater in calibrating the intensity of review than is typically acknowledged. The consequence is that grounds of review on their own have

⁷⁶ See Fordham (note 71 above) at 188. See also D Feldman 'Convention Rights and Substantive Ultra Vires' in C Forsyth (ed) *Judicial Review and the Constitution* (2000) 245, 257.

⁷⁷ M Fordham *Judicial Review* (6th Edition, 2012) 196.

⁷⁸ C Saunders 'Constitution as a Catalyst' (2012) 10 *New Zealand Journal of Public and International Law* 143, 148. For further elucidation that courts are typically allowed to state their own opinion on questions of law, see T Endicott 'Questions of Law' (1998) 114 *Law Quarterly Review* 292; P Craig 'Judicial Review, Appeal and Factual Error' (2004) *Public Law* 788; R Williams 'When is an Error not an Error?' (2007) *Public Law* 793; HWR Wade 'Anglo-American Administrative Law' (1966) 82 *Law Quarterly Review* 226.

limited relevance in the calibration of the intensity of review.⁷⁹ In truth, it is more appropriate to focus our measure of the intensity of review on the order that the court deems necessary to award. It is the court's ruling as to what to do with the decision of the administrator or executive which alerts us as to the intrusiveness of the review. More specifically, we care more about whether a court has made a particular decision in circumstances which should instead have urged the court to show restraint, whether it be under the grounds of lawfulness, reasonableness, irrationality or procedural fairness. The mere fact that the court has asked the executive to proffer argument as to the reasonableness of its decision is not in and of itself dispositive of the intrusiveness of the judiciary. In other words, the intensity of the review is truly manifested in the reaction, and the acceptable range of reasons for that reaction, of the court to those arguments. Does the court accept them as sufficient on the basis of its lack of expertise or democratic legitimacy? Or does the court regard itself fully capable and under a constitutional obligation to stand in the way of a decision made by another branch of government? These are the questions which most directly signify the intrusiveness of the court and the manner in which it is calibrated.

For this picture to become clearer, it may be helpful to have a more detailed breakdown of the role of grounds of review in the review process. Mark Elliott has provided a useful deconstruction of the stages of deference in judicial review which provides a basis on which to picture the role of the various grounds of review.⁸⁰ Elliott begins by pointing out that deference can be broken down into two stages.⁸¹ The first stage he has termed the 'starting-point deference' which denotes the framework within which the adjudication takes place. The second is 'adjudicative deference', which is the deference available to the courts during the adjudicative process.

Starting-point deference is placed within the rubric of a justification thesis.⁸² Elliott argues that the first step a court takes in setting the intensity of review applicable in relation to an impugned decision is to set the bar that the state has to reach in order to justify that decision. This the court does by framing the review in terms of the grounds under which the review will take place. Each ground is accompanied by a standard of justification that is either more or less burdensome on the state in terms of the justification that it must offer. Therefore, asking whether a decision is reasonable is less deferential than asking if the decision is

⁷⁹ I would not go as far as Allan's depiction of grounds of review as lacking in meaning, being empty vessels, which only take tangible form when applied to particular circumstances. See TRS Allan 'Doctrine and Theory in Administrative Law: An Elusive Quest for the Limits of Jurisdiction' (2003) *Public Law* 429. See also R Posner 'What is Obviously Wrong with the Federal Judiciary, Yet Eminently Curable, Part I' (2016) 19 *Green Bag* 187 (A similar critique of standards of review in the United States).

⁸⁰ M Elliott 'From Bifurcation to Calibration: Twin-Track Deference and the Culture of Justification' in M Elliott & H Wilburg (eds) *The Scope and Intensity of Substantive Review* (2015) 61.

⁸¹ See also Price (note 69 above) at 53. (Price makes a similar argument. His argument is, however, less nuanced than that of Elliott. Price seems to make a hard distinction between the variability of review offered by choosing the ground of review, and the variability achievable within the ground of review. Elliott on the other hand takes into account the inseparability of the two forms of variability. He accepts that the variability offered by one can be mediated, enhanced or modified by the variability offered by the other.)

⁸² Elliott (note 80 above) at 66.

rational or lawful. The variation in the standard of review or intensity of review is thus established by choosing to subject the decision to higher or lower standards of justification.

The second stage of the process of calibrating the intensity of the review takes place during the course of adjudication. This stage Elliott titles ‘adjudicative deference’. Here the court must decide whether the state has met the justificatory burden set by the applicable grounds of review. This gives the court the opportunity to take into account all the peculiar circumstances of the case before it. In particular, the court may consider the importance of the fundamental value impacted by the impugned decision or the normative considerations regarding its ability to make a decision with respect to the impugned decision. In other words, the court may place greater weight on the reasons offered by the state if the court believes, for instance, that the state is better placed to understand the factors involved in the decision.

The problem here is that if the level of intrusiveness is calculated only after both the ground of review and the level of applicable deference are taken into account, how then does one compare the level of intrusiveness of grounds of review as an objective fact? Adjudicative deference is a variable contingent on the facts of a particular case. Unless deference is objectively understood to respond differently to each ground of review in a way that ensures that each ground of review is more or less intrusive, then there is no basis to place them in a hierarchy of least to most intrusive forms of review.

Perhaps another way to view the issue is to consider what we typically understand ‘intrusion’ to entail. When the judiciary is accused of an overly-intrusive judgment, the harm is characterised as a separation of powers breach, or, less commonly, an instance of a court grappling with issues beyond its understanding. Given this, we would ordinarily say that a ground of review is unduly intrusive when it has produced either of these harms. More particularly for our purposes we need to determine at what point during the review can we expect to see either of these harms materialise. Plainly, one could not argue that the harm of a separation of powers breach, or a competence concern arises when the court identifies the applicable ground of review. By asking whether a decision was reasonable or rational and having the decision-maker put up justifications for its decision that meet the acceptable standard cannot mean that the court has substituted its decision for that of the decision-maker. It is only when the court attempts to pass judgment on whether the justifications are sufficient in meeting the applicable standard that we are concerned about the court’s expertise, competence or democratic legitimacy. It is when attempting to decide whether the justifications are satisfactory that a court will employ some form of adjudicative deference. Without an *ex ante* understanding of that deference, there is no way to determine whether the court will be intrusive before the case is decided.

What we have shown is that grounds of review do little by themselves to calibrate the intensity of review. As such, they constitute a poor basis for distinguishing between the levels of intensity offered by either legality or PAJA review. A necessary component of such a distinction is the form of deference that a court is compelled to use in a legality or PAJA review. Yet with the un-nuanced understanding of deference that the courts have developed, it may be too early

to label either pathway of review as necessarily being more deferential than the other.

Perhaps, then, the argument should be that the primary function of the grounds of review by themselves is not to play the definitive part in the calibration of the intensity of review, but instead to perform an organisational role in the review process. The grounds of review primarily indicate the questions the court is asking the decision-maker and indicates to the decision-maker what type of justifications it must put forward in order to meet the applicable standard. They perhaps have the incidental effect that greater substance of the decision may fall under the scope of the review when certain grounds are applied as opposed to others. However, the identity of the ground by itself has little effect on the intrusiveness of the review, other than the manner in which the court assesses the justification of the decision-maker.

This issue may, however, be little more than an academic debate. The overlapping grounds of review under the principle of legality and PAJA may make relying on the content of the two pathways for a distinction all but impossible. In reality, the potential for the expansion of the principle of legality is endless. The principle's generality means that it could, in theory, include all the elements that reside in PAJA and perhaps much more.⁸³ The risk of all review collapsing into legality review becomes more than a notional concern with the Court's proclamation of the open-textured nature of the principle of legality.⁸⁴ This is not at all surprising and is maybe not a bad idea.⁸⁵ As the nature of the decisions that are taken on review evolve, there may be a corresponding need for the tools with which a court conducts review to evolve. Despite this, the more obvious concern is this: What are thought to be the more deferential of the grounds of review that exist under the principle of legality, also exist under PAJA review. In other words, the reason that the Court may like to turn to the principle of legality also applies with respect to PAJA review.

If the Court hopes to maintain that a desire to show greater deference is a reason to choose to apply the principle of legality over PAJA review, then greater work needs to be done on the manner in which the principle of legality operates. The Court should aim to ensure that the nominally deferential grounds of review in the principle of legality and those in PAJA review produce two distinct degrees of deference. This can perhaps best be achieved by defining *and* distinguishing the various forms of adjudicative deference that apply with respect to the grounds of review within the principle of legality and PAJA review respectively. A more extreme solution that maintains the distinction between the principle of legality and PAJA review on the basis of a difference in the grounds of review that each

⁸³ C Hoexter 'The Rule of Law and the Principle of Legality in South African Administrative Law Today' in M Carnelley & S Hoctor (eds) *Law, Order and Liberty: Essays in Honour of Tony Mathews* (2011) 55. See also Price (note 44 above) at 656.

⁸⁴ See Sachs J's judgment in *New Clicks* (note 28 above) at para 614 (states that legality 'is an evolving concept in our jurisprudence, whose full creative potential will be developed in a context-driven and incremental manner').

⁸⁵ UK courts have similarly acknowledged that grounds of review are a multiplying breed of tools. See Fordham (note 71 above).

contains is for the Court to end, and possibly reverse, its campaign to make the principle of legality into the more flexible and robust analog of PAJA review.

V THE FUTURE OF THE PRINCIPLE OF LEGALITY

The principle of legality is obviously an important development in the Court's jurisprudence. PAJA both explicitly excludes executive and legislative decisions, as well as other important exercises of public power, from its ambit. Added to this, the legislation presents courts with the difficult task of deciding what constitutes 'administrative action'. The possibility that significant decisions of public authorities are unreviewable is unthinkable, and so the principle of legality fills a potential gap of accountability. Despite its functional use, however, remarkably few attempts have been made to critique the doctrine on its own terms. Much has been written about the doctrine's relationship with PAJA, but few have attempted to explain the borders, mechanics or justifications of this legal instrument developed by the Court.⁸⁶ Here, I only intend to raise a few issues that relate specifically to the question of the principle's standard of review.

The manner in which the Court has approached the review of exercises of public power has meant that when the impugned decision is shown or assumed not to be administrative in nature, the appearance of the principle of legality has been somewhat unpredictable. In fact, it appears that the Court typically raises the principle explicitly only in instances where it is debatable whether the power being exercised is administrative in nature or not. In other words, it only seems to get mention when a central issue before the court is whether PAJA should apply. The lack of predictability is compounded by the Court's relative silence with respect to the reasons for choosing one form of review over another. This pattern of judicial modesty is most apparent in rights-based adjudication where the Court, on the one hand, fails to develop or explain a relationship between PAJA review and review of decisions argued to effect rights.⁸⁷ On the other hand, the Court, after holding that the decision under review is not administrative in nature and therefore does not require the application of PAJA, does not seek to find if the principle of legality should apply.

A recent example is *Shuttleworth*, where the Court dealt with the ministerial decision to impose a 10 per cent levy on wealth transferred abroad. In *Shuttleworth*, Moseneke DCJ laboured through the process of showing that the decision was not administrative in nature and yet made no mention of the principle of legality.⁸⁸

⁸⁶ See Price (note 44 above) at 658 (Raises a fundamental question regarding the justification of the principle of legality). See also Fowkes (note 21 above) at 23.

⁸⁷ S Rose-Ackerman, J Fowkes & S Egidy *Due Process of Lawmaking: The United States, South Africa, Germany and the European Union* (2016) 150. See also *Zealand v Minister of Justice and Constitutional Development and Another* [2008] ZACC 3, 2008 (4) SA 458 (CC), 2008 (6) BCLR 601 (CC), 2008 (2) SACR 1 (CC) (the Court was especially intrusive in the case of a fundamental rights violation).

⁸⁸ See *South African Reserve Bank and Another v Shuttleworth and Another* [2015] ZACC 17, 2015 (5) SA 146 (CC), 2015 (8) BCLR 959 (CC) at para 35. (Moseneke DCJ relies on *Permanent Secretary of the Department of Education and Welfare, Eastern Cape Province and Another v Ed-U-College* [2000] ZACC 23, 2001 (2) SA 1 (CC), 2001 (2) BCLR 118 (CC) to show that the decision was one of high policy and therefore did not constitute an administrative decision.)

The result is the seemingly haphazard use of a broad and general principle and the unpredictable application of a standard of review to any given case.⁸⁹

It is clear that the principle of legality as an abstract value is expected of all exercises of public power. What *Motau* has pointed out is that the principle is used directly by courts with respect to some categories of powers, while with others the courts adopt some form of proxy that achieves the same aim.⁹⁰ The problem remains that there are cases on the fringes of these categories that receive uneven application of the principle of legality. This uneven distribution of the principle of legality has been pointed out before. However, *Motau* reveals the tension that underlies the broad application of the principle of legality. Quite simply, in instances where the principle is used, ignored, or simply not mentioned, there is no clear distinction between the standards of review that apply to these categories of decisions.

However, developing a coherent doctrine of use for the principle of legality with its attendant standard(s) of review is not an easy task. In order to accomplish this, the Court must take into account that the principle applies to the exercise of all public power and principally applies to decisions of the executive and legislative branches of government. The trick will be to develop standards of review and therefore methods of deference that are responsive to the nature of power under review as well as peculiar circumstances that arise in any review, without expanding the width of variability to the point of arbitrariness.⁹¹

What complicates this task is that these are very broad categories and contain institutions of various types that exercise powers of various natures. As discussed above, standards of review, and our desire to show deference, to a large extent, is dependent on principles of separation of powers and institutional competence. Under the umbrella of these principles, one could pack a great number of norms that would form the justification for a court's demeanour towards a particular decision. For instance, the separation of powers guards liberty against a tyrannical state, and enhances democratic accountability.⁹² Matching tasks according to institutional competence may enhance democratic accountability by publically and transparently ensuring that a decision is taken by the institution most capable of producing the superior decision.⁹³ The complication arises as any governmental institution which possesses powers or functions that are either executive or legislative in nature could inhabit a place within the constitutional scheme which would have different implications with respect to any of these

⁸⁹ Further examples include *MEC for Social Development, Western Cape and Others v Justice Alliance of South Africa and Another* [2016] ZASCA 88 (the SCA bases its decision squarely on the separation of powers and does not refer to the principle of legality) and *International Trade Administration Commission v SCAP South Africa (Pty) Ltd* [2010] ZACC 6, 2012 (4) SA 618 (CC), 2010 (5) BCLR 457 (CC).

⁹⁰ See also H Corder 'Reviewing Executive Action' in J Klaaren (ed) *A Delicate Balance, The Place of the Judiciary in a Constitutional Democracy* (2006) 73, 76.

⁹¹ See M du Plessis & S Scott 'The Variable Standard of Rationality Review: Suggestions for Improved Legality Jurisprudence' (2013) 130 *South African Law Journal* 597.

⁹² See A Huq 'Libertarian Separation of Powers' (2014) 8 *New York University Journal of Law and Liberty* 1006. See also J Waldron *Political Theory* (2016) 63. (Waldron has argued that separation of powers furthers the rule of law. By slowing down the process of decision-making, separation of powers promotes regularity and stability.)

⁹³ A Huq 'The Institution Matching Canon' (2015) 106 *Northwestern University Law Review* 417.

norms. For instance, an executive agency could be under direct control of a ministerial department and thus more democratically accountable. On the other hand, the same agency could be independent of any executive control and thus be less accountable. Review in each instance would be calibrated differently in order to account for this distinction.

The complications are not limited to identifying the plurality of norms that drive the formulation of the standard of review. These norms can also conflict with one another. For instance, in the example above, the independent executive agency could be applying its executive powers to highly complex issues. On the one hand, the court is faced with an unaccountable institution which may encourage it to apply a stricter standard of review. Yet on the other hand, its lack of competence may necessitate a more lenient approach. Standards of review that comprise the principle of legality will have to take this into account.

With the principle of legality, the problem is compounded with the application of the principle to a far broader array of institutions than those within the traditional conception of ‘public bodies’. In *AAA Investments*, the Court accepted that any institution, whether private or public in origin, that performs a public function is considered to be an ‘organ of state’ under s 239 of the Constitution.⁹⁴ The decisions of these institutions, if also considered executive in nature, are then subject to review under the principle of legality. This opens the door to a substantial variety of institutions that exhibit an inexhaustible number of characteristics relevant for the application of a standard of review.

It may be that the Court would prefer that the principle of legality remain broad and obscure, and its application remain unpredictable. Such latitude allows the Court to achieve other aims, such as the preservation of its own legitimacy or the flexibility to respond to peculiar facts of individual cases. Courts may, within the space created by the principle of legality, whether explicitly used or not, operate based on motives removed from the strict doctrinal considerations of law. This doctrinal vacuum would allow for the more pragmatic style of decision-making that some scholars have noted.⁹⁵ Despite being faced with the task of determining constitutional meaning, courts focus on their appropriate or strategic role in the adjudicatory process.⁹⁶ If *Motau* is read to further the objective of increasing judicial latitude, then the pleasant aspect of its reasoning could be that, in clearly pointing out the pragmatic use of the principle of legality and the notion of deference, the Court was being more transparent than it usually is.⁹⁷

⁹⁴ See *AAA Investments v Micro Finance Regulatory Council and Another* [2006] ZACC 9, 2007 (1) SA 343 (CC), 2006 (11) BCLR 1255 (CC) at para 30. This was confirmed in *Allpay Consolidated Investment Holdings (Pty) Ltd and Others v Chief Executive Officer of the South African Social Security Agency and Others* [2014] ZACC 12, 2014 (4) SA 179 (CC), 2014 (6) BCLR 641 (CC) at para 53. See also M Finn ‘Allpay Remedy: Dissecting the Constitutional Court’s Approach to Organs of State’ (2015) 6 *Constitutional Court Review* 258.

⁹⁵ See, eg, T Roux *The Politics of Principle: The First South African Constitutional Court, 1995–2005* (2013).

⁹⁶ See also most recently D Strauss ‘The Supreme Court, Foreword: Does the Constitution Mean What It Says’ (2015) 129 *Harvard Law Review* 1 (Argues that courts ought to, and do, take into account extraneous factors in order to determine what its appropriate role is as opposed to a process which focuses entirely on the constitutional text).

⁹⁷ Something that would be welcomed by some. See S Woolman ‘The Amazing, Vanishing Bill of Rights’ (2007) 124 *South African Law Journal* 762.

On the other hand, the Court in *Motau* could have made a valuable initial step in making what is a complicated bifurcation of administrative law more workable. Simply, if we know what each pathway of review is for and when to use them, we will have a simpler review process.

VI CONCLUSION

What has become quite clear is that there needs to a better dividing line between the principle of legality and PAJA review. If ever it was suitable to point to the distinct grounds of review each contains in order to tell them apart, it is difficult to see how that can be justified with current case law. The same can be said of the use of any notion of deference that the Court may currently have. Yet the Court in *Motau* has quite clearly claimed that the two pathways of review can be distinguished based on their distinct standards of review. On the basis of this claim, the belief is that the principle of legality is the more appropriate pathway of review in cases dealing with executive decisions. This is so important a distinction that the Court argues that it should play a role in the determination of what is or is not executive action. The overall conclusion is that the principle of legality, in the Court's mind, should be thought of as a less invasive form of review than administrative law review and used when the court feels that deference is owed to the decision-maker.

These findings create a puzzle that I think should be, at least partially, the focus of the Court's efforts in the development of the judicial review of exercises of public power. This note, I hope, points towards a few issues that should be of concern. The primary point is that standards of review are made up of two components which have a complicated relationship with one another: the questions being asked of the impugned decision, and the degree of receptivity of the court to any answers that the state puts forward. Both should be clearly articulated in a set of review mechanisms if we are to assume that one is more or less deferential than the other. As a corollary step, courts may have to reconsider the expansion of the principle of the legality which increasingly looks like PAJA review and, instead, focus on finding the distinguishing features of the two forms of review.

The Test for ‘Exceptional Circumstances’ Where an Order of Substitution is Sought: An Analysis of *Trencon* Against the Backdrop of the Separation of Powers

Lauren Kohn*

I INTRODUCTION

There is perhaps some irony in the fact that a doctrine not mentioned by name in our Constitution¹ – the separation of powers – has been, and continues to be, a hot talking point. One of the areas in which it generates debate is that of the courts’ remedial powers, particularly in constitutional matters. Section 172(1)(b) of the Constitution vests the courts with a generous discretionary power to make ‘any order that is just and equitable’ in constitutional matters. Acting under the rubric of these somewhat amorphous guiding tenets of justice and equity, our courts are required to do some careful balancing: balancing of the need to ensure both a degree of certainty and a degree of flexibility when carving out the requisites of an appropriate remedy; balancing of the need to ensure they fulfil their role as guardians of the Constitution by awarding effective relief where rights need to be vindicated, while at the same time remaining conscious of the need not to overstep into the ‘boggy terrain’² of policy which belongs in the legislative, executive and, to a degree, administrative heartlands. This balancing act is not an easy one. It is particularly tricky where ‘exceptional circumstances’ call for exceptional relief.

Section 8(1)(c)(ii)(aa) of the Promotion of Administrative Justice Act 3 of 2000 (PAJA), the constitutionally mandated legislation that seeks to give effect to the s 33 rights to ‘just administrative action’, empowers courts in judicial review proceedings to make a ‘just and equitable’ order ‘substituting or varying the administrative action or correcting a defect resulting from the administrative

* Lecturer, Faculty of Law, University of Cape Town, attorney of the High Court of South Africa and legal consultant at Caveat Legal. I am grateful to my colleagues, Professor Hugh Corder and Associate Professor Alistair Price, for the useful discussions we had regarding *Trencon*, and for their valued mentorship. I must also thank my friend and colleague, Raisa Cachalia, for our interesting conversations regarding the complexities of *Trencon*.

¹ Constitution of the Republic of South Africa, 1996.

² *Paulsen and Another v Slip Knot Investments 777 (Pty) Ltd* [2015] ZACC 5, 2015 (3) SA 479 (CC) at para 58.

action’ in ‘exceptional cases’. The PAJA, however, fails to provide legislative guidance as to what exceptional circumstances might call for this exceptional remedy. This is thus an area where ‘[t]he common law informs the provisions of PAJA’.³ Out of the common law, several fairly loosely conceived factors in this exceptional circumstances test have crystallised. However, their substantive content, interplay and pecking order in the enquiry have been fairly unclear. The recent judgment of the Constitutional Court in *Trencon Construction (Pty) Ltd v Industrial Development Corporation of South Africa Ltd* aims to ‘clarify the test for exceptional circumstances where a substitution order is sought’⁴ and seems to go some way in doing so. In particular, the Court picks up where Plasket J left off in *Intertrade Two (Pty) Ltd v MEC, Roads & Public Works, Eastern Cape*, where he stated that ‘[t]he availability of proper and adequate information and the *institutional competence* of the court to take the decision for the administrative decision-maker are *necessary prerequisites* that *must* be present, *apart from* “exceptional circumstances”’, before a court can legitimately assume an administrative decision-making function’.⁵

Informed by the degree of deference required by the separation of powers in awarding appropriate relief, the *Trencon* Court puts centre-stage those factors that go to the institutional competence of the courts and, at the same time, emphasises the vital overarching role of the notion of fairness in the enquiry. In doing so, the judgment – although rather unclear in parts – goes some way to achieving that sensitive balance between the need for a degree of both certainty and flexibility, and the need to avoid both judicial timidity and judicial excess in the granting of remedies. On the whole, it is both a principled and pragmatic judgment showing mindfulness of the vital role public procurement plays in our society and concomitantly recognising the dangers of the abuse, or merely the poor exercise, of public power in this context.

In this comment I discuss *Trencon* against the backdrop of the separation of powers and the ‘formal and flexible rules of restraint’⁶ that flow from it and ought to guide the courts in exercising their remedial powers, particularly those of an exceptional nature such as substitution. In doing so I aim, in particular, to interpret the Court’s formulation of the exceptional circumstances test in a constructive and accessible manner, for this formulation will no doubt serve as the litmus test for the courts in future and will play a significant practical role when an aggrieved party is deciding whether to litigate, particularly in the procurement context.

³ *Bato Star Fishing (Pty) Ltd v Minister of Environmental Affairs and Tourism and Others* [2004] ZACC 15, 2004 (4) SA 490 (CC), 2004 (7) BCLR 687 (CC) (‘*Bato Star*’) at para 22. On the relationship between the common law, the PAJA and s 33 of the Constitution, see L Kohn ‘Our Curious Administrative Law Love Triangle: The Complex Interplay Between the PAJA, the Constitution and the Common Law’ (2013) 28 *South African Public Law* 22.

⁴ *Trencon Construction (Pty) Ltd v Industrial Development Corporation of South Africa Ltd and Another* [2015] ZACC 22, 2015 (5) SA 245 (CC), 2015 (10) BCLR 1199 (CC) (‘*Trencon*’) at para 32.

⁵ *Intertrade Two (Pty) Ltd v MEC for Roads and Public Works, Eastern Cape, and Another* [2007] ZAECHC 149, 2007 (6) SA 442 (Ck), [2008] 1 All SA 142 (Ck) (‘*Intertrade*’) at para 43 (emphasis added).

⁶ L Kohn ‘The Burgeoning Constitutional Requirement of Rationality and the Separation of Powers: Has Rationality Review Gone Too Far?’ (2013) 130 *South African Law Journal* 810, 820 (‘Rationality Review’).

II THE SEPARATION OF POWERS AS A GUIDE TO THE COURTS IN THE EXERCISE OF THEIR REMEDIAL POWERS

Although not mentioned by name, it is axiomatic that by its design our Constitution implicitly entrenches the doctrine of the separation of powers.⁷ In theory the doctrine comprises two fairly straightforward propositions.⁸ The first is that to prevent the abuse of public power, it must not be concentrated in any one arm of state but must rather be divided amongst them.⁹ The second proposition qualifies this first one: it is the principle of checks and balances pursuant to which the separation is not absolute insofar as each of the three branches exercises some form of 'check' over the power of the others.¹⁰ As I have noted elsewhere, '[t]he judiciary provides the most crucial check against abuses of state power, and this is most obviously done through the "potentially awesome power" of judicial review'.¹¹ Within the separation of powers, the judiciary is thus both player and referee and therefore has the role of policing compliance with the Constitution by the other arms of state (as well as the administration),¹² while determining for itself just how far to go in exercising this policing function. Given this sensitive dual role demanded by the separation of powers, a 'delicate balancing'¹³ is required in the discharge of the judicial function. Thus, in *ITAC v SCAW South Africa (Pty) Ltd*, Moseneke DCJ noted the following:

Where the Constitution or valid legislation has entrusted specific powers and functions to a particular branch of government, courts may not usurp that power or function by

⁷ See *South African Association of Personal Injury Lawyers v Heath and Others* [2000] ZACC 22, 2001 (1) SA 883 (CC), 2001 (1) BCLR 77 (CC) at para 19. On the doctrine generally see, eg, K O'Regan 'Checks and Balances: Reflections on the Development of the Doctrine of Separation of Powers under the South African Constitution' (2005) 8 *Potchefstroom Electronic Law Journal* 120; P Labuschagne 'The Doctrine of Separation of Powers and its Application in South Africa' (2004) 23 *Politeia* 84; PN Langa 'The Separation of Powers in the South African Constitution' (2006) 22 *South African Journal on Human Rights* 2; S Ngcobo 'South Africa's Transformative Constitution: Towards an Appropriate Doctrine of Separation of Powers' (2011) 22 *Stellenbosch Law Review* 37; and GE Devenish 'The Doctrine of Separation of Powers with Special Reference to Events in South Africa and Zimbabwe' (2003) 66 *Tydskrif vir Hedendaagse Romeins-Hollandse Reg* 84.

⁸ On the two propositions, see *Ex parte Chairperson of the Constitutional Assembly: In re Certification of the Constitution of the Republic of South Africa, 1996* [1996] ZACC 26, 1996 (4) SA 744 (CC), 1996 (10) BCLR 1253 (CC) at para 109: 'The principle of separation of powers, on the one hand, recognises the functional independence of branches of government. On the other hand, the principle of checks and balances focuses on the desirability of ensuring that the constitutional order, as a totality, prevents the branches of government from usurping power from one another. In this sense it anticipates the necessary or unavoidable intrusion of one branch on the terrain of another.'

⁹ Thus, s 43 of the Constitution entrusts legislative authority to the legislatures at national, provincial and local government level, s 85 vests the executive authority of the Republic at national level in the President and his cabinet, and s 165(1) vests the judicial authority in the courts.

¹⁰ An example of an operational provision that epitomises a 'check' within the system of checks and balances is s 172 of the Constitution.

¹¹ Kohn 'Rationality Review' (note 6 above) at 816.

¹² See *Glenister v President of the Republic of South Africa and Others* [2008] ZACC 19, 2009 (1) SA 287 (CC), 2009 (2) BCLR 136 (CC) at para 33, where the Court notes that '[i]t is a necessary component of the doctrine of separation of powers that courts have a constitutional obligation to ensure that the exercise of power by other branches of government occurs within constitutional bounds. But even in these circumstances, courts must observe the limits of their powers.'

¹³ *De Lange v Smuts NO and Others* [1998] ZACC 6, 1998 (3) SA 785 (CC), 1998 (7) BCLR 779 (CC) at para 60.

making a decision of their preference. That would frustrate the balance of power implied in the principle of separation of powers. The primary responsibility of a court is not to make decisions reserved for or within the domain of other branches of government, but rather to ensure that the concerned branches of government exercise their authority within the bounds of the Constitution. This would especially be so where the decision in issue is policy-laden as well as polycentric.¹⁴

This warning rings equally true when it comes to the functions of administrative agencies which ‘the court[s] should take care not to usurp’.¹⁵ To achieve this delicate balancing demanded by the separation of powers, and thereby ensure that the ‘defensible limits of judicial review’¹⁶ are observed, what I have explained to be ‘formal and flexible rules of restraint’¹⁷ serve to guide the courts in exercising their review function and awarding appropriate relief. A ‘significant’¹⁸ formal limit is the review / appeal dichotomy pursuant to which a court exercising review jurisdiction over the legality of a decision-making process should be mindful of the need not to slip into an appeal-style assessment of the correctness or otherwise of the outcome in a given case.¹⁹ Thus, in the recent case of *City of Cape Town v South African National Roads Agency Ltd*, Binns-Ward and Boqwana JJ noted that:

[a]ppeals entail reconsidering the merits of an impugned decision (a rehearing in effect), with the appellate tribunal being empowered to substitute its decision for that of the first instance decision-maker. Reviews, on the other hand, are not concerned, other than sometimes incidentally, with the merits ... and only exceptionally will they give rise to a substitutive decision.²⁰

Section 8(1)(c)(ii)(aa) of the PAJA contemplates such an exceptional situation. The power to substitute must therefore be exercised judiciously and in accordance with the requisite degree of deference or ‘respect’²¹ (the flexible self-imposed rule of restraint)²² called for by the facts of a given case. In *Bato Star* our Constitutional Court endorsed Hoexter’s account of judicial deference as:

[A] judicial willingness to appreciate the legitimate and constitutionally-ordained province of administrative agencies; to admit the expertise of those agencies in policy-laden or polycentric issues, to accord their interpretations of facts and law due respect; and to be sensitive in general to the interests legitimately pursued by administrative bodies and the practical and financial constraints under which they operate. This type of deference is perfectly consistent with a concern for individual rights and a refusal to tolerate corruption and maladministration. *It ought to be shaped ... by a careful weighing up of the need for*

¹⁴ *International Trade Administration Commission v SCAW South Africa (Pty) Ltd* [2010] ZACC 6, 2012 (4) SA 618 (CC), 2010 (5) BCLR 457 (CC) at para 95.

¹⁵ *Bato Star* (note 3 above) at para 45.

¹⁶ P Lenta ‘Judicial Restraint and Overreach’ (2004) 20 *South African Journal of Human Rights* 544.

¹⁷ Kohn ‘Rationality Review’ (note 6 above) at 820.

¹⁸ *Bato Star* (note 3 above) at para 45.

¹⁹ See Kohn ‘Rationality Review’ (note 6 above) at 820.

²⁰ *City of Cape Town v South African National Roads Agency Ltd and Others* [2015] ZAWCHC 135, 2015 (6) SA 535 (WCC), 2016 (1) BCLR 49 (WCC) at para 8.

²¹ *Bato Star* (note 3 above) at para 46.

²² See my discussion of this notion of deference in Kohn ‘Rationality Review’ (note 6 above) at 822–824.

– and the consequences of – judicial intervention. Above all, it ought to be shaped by a conscious determination not to usurp the functions of administrative agencies; not to cross over from review to appeal.²³

This form of deference thus requires an honest assessment by a court of its institutional competence in a particular case. Context matters greatly, for '[d]etermining the boundaries of the courts' proper role ... cannot be reduced to a simple test or formula; it will vary according to ... the context of each case'.²⁴ Implicit in this recognition of the need for a degree of flexibility is a concomitant recognition of the flipside to the coin of deference: certain circumstances may call for less deference, more searching review and/or a more robust approach to remedy – such as substitution. This is consonant with the principle of checks and balances demanded by the separation of powers which must be understood as “operationally defined” by the Constitution.²⁵

Section 172 of the Constitution is an apt example of a vital ‘operational check’ within the separation of powers: it grants the courts a wide discretion to ‘make any order that is just and equitable’ in fulfilling their mandate to declare invalid law or conduct that is inconsistent with the Constitution.²⁶ This phrase is mirrored in s 8(1) of the PAJA – legislation that enjoys a special dual status insofar as it is constitutionally mandated *and* a product of the exercise of the democratic will as expressed by the legislature – which, in an open list of possible remedies, expressly contemplates substitution as a just and equitable remedy where the circumstances are ‘exceptional’ and thus demand as much. This remedy must therefore be understood as a manifestation of what the separation of powers may in fact require in a particular context. Thus in *Allpay 2* Froneman J noted that

[t]here can be no doubt that the separation of powers attributes responsibility to the courts for ensuring that unconstitutional conduct is declared invalid and that constitutionally mandated remedies are afforded for violations of the Constitution. This means that the Court *must* provide effective relief for infringements of constitutional rights.²⁷

Substitution, although the exception rather than the norm, may be the effective and thus ‘appropriate’²⁸ relief required by the particular facts and thus what justice

²³ *Bato Star* (note 3 above) at para 46, citing C Hoexter ‘The Future of Judicial Review in South African Administrative Law’ (2000) 117 *South African Law Journal* 484, 501–502 (emphasis added).

²⁴ McLachlin J in *RJR-MacDonald Inc v Canada (Attorney-General)* (1995) 31 CRR (2nd) 189, [1995] 3 SCR 199 at para 136.

²⁵ *Dodo v S* [2001] ZACC 16, 2001 (3) SA 382 (CC), 2001 (5) BCLR 423 (CC) at para 17, citing Tribe’s remarks in relation to the US Constitution.

²⁶ Thus Froneman J stated in *AllPay Consolidated Investment Holdings (Pty) Ltd and Others v Chief Executive Officer of the South African Social Security Agency and Others (No 2)* [2014] ZACC 12, 2014 (4) SA 179 (CC), 2014 (6) BCLR 641 (CC) (*AllPay 2*) at para 45: ‘[T]he answer to the separation-of-powers argument lies in the express provisions of s 172(1) of the Constitution. The corrective principle embodied there allows correction to the extent of the constitutional inconsistency.’

²⁷ *Ibid* at para 42 (emphasis added).

²⁸ In *Fose v Minister of Safety and Security* [1997] ZACC 6, 1997 (3) SA 786 (CC), 1997 (7) BCLR 851 at para 69 the Court held that ‘an appropriate remedy must mean an effective remedy’.

and equity in a given case necessitate.²⁹ Accordingly, the test for exceptional circumstances must be understood against the backdrop of the first principles that the Court has elucidated when it comes to the question of remedy. These were crisply summarised by Moseneke DCJ in *Steenkamp NO v Provincial Tender Board, Eastern Cape*:

It goes without saying that every improper performance of an administrative function would implicate the Constitution and entitle the aggrieved party to appropriate relief. In each case the remedy must fit the injury. The remedy *must be fair* to those affected by it and yet vindicate effectively the right violated. It must be *just and equitable in the light of the facts, the implicated constitutional principles*, if any, and the controlling law... The purpose of a public law remedy is to pre-empt or *correct* or reverse an improper administrative function. In some instances the remedy takes the form of an order to make or not to make a particular decision Ultimately the purpose of a public law remedy is to afford the prejudiced party administrative justice, to advance efficient and effective public administration compelled by constitutional precepts and at a broader level, to entrench the rule of law.

...

Examples of public remedies suited to vindicate breaches of administrative justice are to be found in s 8 of the PAJA ... [which] confers on a court in proceedings for judicial review a generous jurisdiction to make orders that are ‘just and equitable’.³⁰

The overarching question when it comes to the remedy of substitution is thus whether justice and equity – code for simple fairness – demand as much on the distinctive facts. This is nothing novel: the guiding consideration of fairness was emphasised even in the pre-PAJA common-law jurisprudence.³¹ However, given the extraordinary nature of the remedy in the context of the separation of powers and the courts’ *sui generis* institutional competence, unfettered flexibility is both unhelpful and unwarranted. For this reason, the courts have carved out factors to guide the exercise of their discretion and thereby assist in ensuring equilibrium between the need for a degree of both clarity *and* flexibility. Before illustrating how the *Trencon* judgment purports to clarify this exceptional circumstances test and to ensure this equilibrium, I turn briefly to consider this test’s jurisprudential roots and the legal principles it entails.

III WHEN TO SUBSTITUTE: THE COMMON-LAW GUIDELINES

Given that the PAJA fails to provide any legislative guidance as to when a case will be ‘exceptional’ and thus call for substitution, the common-law principles continue to be instructive. It is a well-established principle of our common law ‘that the courts will be reluctant to substitute their decision for that of the original

²⁹ See K Roach & G Budlender ‘Mandatory Relief and Supervisory Jurisdiction: When is it Appropriate, Just and Equitable?’ (2005) 122 *South African Law Journal* 325, 351, where the authors highlight that ‘[d]ifferent remedial routes may be appropriate in different circumstances, but the ultimate destination that the courts should insist upon is compliance with the constitution. In the final analysis, the test is one of effectiveness. Court orders that are not effective undermine respect for the courts, for the rule of law, and for the constitution itself.’

³⁰ *Steenkamp NO v Provincial Tender Board of the Eastern Cape* [2006] ZACC 16, 2007 (3) SA 121 (CC), 2007 (3) BCLR 300 (CC) at paras 29–30 (emphasis added).

³¹ See the discussion at III below.

decision maker'.³² This necessary reluctance to intervene and substitute flows directly from the separation of powers, which requires the courts to recognise their institutional limitations and respect the comparative institutional competence of administrative agencies, particularly in polycentric matters. Thus, in *Gauteng Gambling Board v Silverstar Development Ltd* Heher JA noted that '[a]n administrative functionary ... is generally best equipped by the variety of its composition, by experience, and its access to sources of relevant information and expertise to make the right decision. The court typically has none of these advantages and is required to recognise its own limitations'.³³ Remittal is thus the general rule and 'almost always the prudent and proper course'.³⁴ However, sometimes rules need to be broken, and our courts have long since recognised that there may be unusual instances in which substitution is the appropriate remedy. On a big-picture level, this determination has essentially always required the courts to ask the question whether 'a decision to exercise a power should not be left to the designated functionary'.³⁵

Established principles have emerged from the common law to guide the courts in answering this question. In particular, certain factors have crystallised out of the cases so as to afford the enquiry (and thus prospective litigants) a degree of certainty.³⁶ However, these factors have never been considered in the abstract: the cases show that '[f]airness to both sides has always been and will almost certainly remain an important consideration',³⁷ which may be decisive in tipping the scales. Thus, in the early case of *Livestock and Meat Industries Control Board v Garda*, the court emphasised that it has 'a discretion, to be exercised judicially upon a consideration of the facts of each case, and ... although the matter will be sent back if there is no reason for not doing so, *in essence it is a question of fairness to both sides*'.³⁸ This principle permeates the post-PAJA jurisprudence too, an arguably unsurprising fact given the constitutional imprimatur of the need to ensure 'just and equitable' relief. For example, in *Commissioner, Competition Commission v General Council of the Bar of South Africa* Hefer AP held that '[a]ll that can be said is that considerations of fairness may in a given case require the court to make the decision itself provided it is able to do so'.³⁹ When might a court be able to do so? The following open list of considerations, or factors, has emerged from the pre- and post-PAJA case law and may, on the facts of a given case, prompt a decision to substitute.⁴⁰

³² *Premier, Mpumalanga and Another v Executive Committee, Association of State-Aided Schools, Eastern Transvaal* [1998] ZACC 20, 1999 (2) SA 91 (CC), 1999 (2) BCLR 151 (CC) at para 50.

³³ *Gauteng Gambling Board v Silverstar Development Ltd and Others* [2005] ZASCA 19, 2005 (4) SA 67 (SCA) ('*Silverstar*') at para 29.

³⁴ *Ibid.*

³⁵ *Ibid* at para 28.

³⁶ On these factors, see the discussion in section IV.A below.

³⁷ C Hoexter *Administrative Law in South Africa* (2nd Edition, 2012) 553.

³⁸ *Livestock and Meat Industries Control Board v Garda* 1961 (1) SA 342 (A) at 349G (emphasis added).

³⁹ *Commissioner, Competition Commission v General Council of the Bar of South Africa and Others* [2002] ZASCA 101, 2002 (6) SA 606 (SCA) ('*Competition Commission*') at para 15.

⁴⁰ For a helpful summary, see L Baxter *Administrative Law* (1984) 681 et seq; *University of the Western Cape and Others v Member of Executive Committee for Health and Social Services and Others* 1998 (3) SA 124 (C) ('*UWC*') at 131D–J; and *Ruyobezu and Another v Minister of Home Affairs and Others* 2003 (5) SA 51 (C), 2003 (8) BCLR 920 (C) ('*Ruyobezu*') at 64G.

- (a) where the court is ‘in as good a position’ and thus as well qualified as the original authority to make the decision;⁴¹
- (b) where ‘the end result is in any event a foregone conclusion and it would merely be a waste of time to order the tribunal or functionary to reconsider the matter’;⁴²
- (c) where additional delay would cause unjustifiable prejudice;⁴³ and
- (d) where ‘the functionary or tribunal has exhibited bias or incompetence to such a degree that it would be unfair to require the applicant to submit to the same jurisdiction again’.⁴⁴

The latter two factors are essentially equitable considerations, while the first two necessitate an assessment of the comparative institutional competence of the court vis-à-vis that of the administrator in the future, should the matter be remitted. While the Court in *Trencon* notes that both pre- and post-PAJA case law seems to suggest that ‘if any factor is established on its own, it would be sufficient to justify an order of substitution’,⁴⁵ it appears to me that those factors that go to institutional competence have in fact previously been considered (even if merely impliedly so) more significant in the enquiry. For example, in the early benchmark case of *JCC*, despite recognising the issue of delay and the fact that ‘a few weeks [could] be saved if the Court assume[d] the Administrator’s function’, Hiemstra J nonetheless went on to note that ‘in addition, the result *must* be a foregone conclusion’.⁴⁶ Plasket J put it even more explicitly in *Intertrade*: ‘The availability of proper and adequate information and the *institutional competence* of the court to take the decision for the administrative decision-maker are *necessary prerequisites* that must be present, *apart from* “exceptional circumstances”, before a court can legitimately assume an administrative decision-making function.’⁴⁷ More recently, in *M v Minister of Home Affairs*, the court canvassed the various factors and concluded by remarking that ‘[o]f course, the court *should* be in a *position*

⁴¹ This consideration was explicitly added by the SCA in *Silverstar* (note 33 above) at para 39 and had also been decisive in earlier cases such as *Theron en Andere v Ring van Wellington van die NG Sendingskerk in Suid-Afrika en Andere* 1976 (2) SA 1 (A) (‘*Theron*’) where, given the judicial nature of the decision in question, the court felt it was fully qualified to substitute. See also *M v Minister of Home Affairs* [2014] ZAGPPHC 649 (‘*M*’) at para 177, *UWC* (note 40 above) at 131G–H and *Competition Commission* (note 39 above) at para 15.

⁴² This common-law principle was stated by Hiemstra J in *Johannesburg City Council v Administrator, Transvaal and Another* 1969 (2) SA 72 (T) (‘*JCC*’) at 76E–F. See also, eg, *Hangklip Environmental Action Group v Minister for Agriculture, Environmental Affairs and Development Planning, Western Cape and Others* [2007] ZAWCHC 41, 2007 (6) SA 65 (C) and *Silverstar* (note 33 above) at paras 38–39.

⁴³ Hoexter (note 37 above) notes at 553 fn 278 that ‘this factor, while it was treated as adding weight to the “foregone conclusion” factor in the *Johannesburg City Council* case, ... has since been regarded as an independent consideration’. See, eg, *Ruyobezza* (note 40 above) at para 49; *M* (note 41 above) at para 175; *ICS Pension Fund v Sithole and Others NNO* [2009] ZAGPHC 6, 2010 (3) SA 419 (T) at para 97; and *Head, Western Cape Education Department and Others v Governing Body, Point High School and Others* [2008] ZASCA 48, 2008 (5) SA 18 (SCA) at para 17.

⁴⁴ *UWC* (note 40 above) at 131D–E. See also, eg, *Mlokoti v Amathole District Municipality and Another* [2008] ZAECHC 184, 2009 (6) SA 354 (ECD) at 380I–381B; *Oskil Properties (Pty) Ltd v Chairman of the Rent Control Board and Others* 1985 (2) SA 234 (SE) at 247E; *Tantoush v Refugee Appeal Board and Others* [2007] ZAGPHC 191, 2008 (1) SA 232 (T) at para 127; and *M* (note 41 above) at paras 170–174.

⁴⁵ *Trencon* (note 4 above) at para 39, and see the cases cited in fn 33.

⁴⁶ *JCC* (note 42 above) at 179 (emphasis added).

⁴⁷ *Intertrade* (note 5 above) (emphasis added).

practically to make the decision';⁴⁸ and in *Justice Alliance of South Africa v Mncube NO* Bozalek J was emphatic in saying that '[m]ost importantly ... , I do not consider that this is a case where the end result is a foregone conclusion'.⁴⁹

The 'clarification' provided by *Trencon* thus seems to be the inevitable outcome of a trend emerging in the case law with its roots in the early pre-PAJA jurisprudence. This trend has, however, not been a clear and uniform one: as the Court notes, '[some of the] earlier case law seemed to suggest that each factor ... may be sufficient on its own to justify substitution. However, it is unclear from more recent case law whether these considerations are cumulative or discrete'.⁵⁰ The Court goes on to cite several recent cases in support of this observation.⁵¹ For example, in *Reizis* the court appears to consider each of the four factors cumulatively in a kind of mechanical tick-box exercise.⁵² The *Trencon* Court thus considers it to be 'of great import that the test for exceptional circumstances be revisited'.⁵³ There is indeed value in the highest court's laying down the tracks of this test and providing greater structure to the enquiry; especially given the inevitable separation-of-powers tensions it creates. I turn now to consider how the *Trencon* judgment seeks to achieve this.

IV *TRENCON*: AN ATTEMPT TO BALANCE CERTAINTY AND FLEXIBILITY IN THE TEST FOR EXCEPTIONAL CIRCUMSTANCES

The facts of *Trencon* emerge out of a procurement dispute, and it is through the prism of both fairness broadly and the regulatory framework governing public procurement that the Court views the issues in this case.⁵⁴ Thus in the introductory paragraph Khampepe J, writing for a unanimous Court, highlights the following:

[T]endering plays a vital role in the delivery of goods and services. Large sums of public money are poured into the process and government wields massive public power when choosing to award a tender. It is for this reason that the Constitution obliges organs of state to ensure that a procurement process is fair, equitable, transparent, competitive and cost-effective. *Where the procurement process is shown not to be so, courts have the power to intervene.*⁵⁵

The Court thus sets the tone for what is to come in the judgment: an implicit highlighting of the vital role that the courts play as watchdogs to address

⁴⁸ *M* (note 41 above) at para 166 (emphasis added).

⁴⁹ *Justice Alliance of South Africa v Mncube NO and Others; In re: Cause for Justice and Another v Independent Communications Authority of South Africa and Others; In re: Doctors for Life International WC v Independent Communications Authority of South Africa and Others* [2014] ZAWCHC 162, 2015 (1) All SA 181 (WCC) at para 187 (emphasis added).

⁵⁰ *Trencon* (note 4 above) at para 46.

⁵¹ *Ibid*, referring to *Radjabu v Chairperson of the Standing Committee for Refugee Affairs and Others* [2014] ZAWCHC 134, 2015 (1) All SA 100 (WCC) at paras 33–39; *Media 24 Holdings (Pty) Ltd v Chairman of the Appeals Board of the Press Council of South Africa and Another* [2014] ZAGPJHC 194 at para 25; *Nucon Roads and Civils (Pty) Ltd v MEC for Department of Public Works, Roads and Transport: NW Province and Others* [2014] ZANWHC 19 at paras 32, 41 and 44; and *Reizis NO v MEC for the Department of Sport, Arts, Culture and Recreation and Others* [2013] ZAFSHC 20 ('*Reizis*') at paras 35–39.

⁵² *Reizis* (note 51 above) at paras 35–39.

⁵³ *Trencon* (note 4 above) at para 41.

⁵⁴ *Ibid*: see, eg, at paras 75 and 78.

⁵⁵ *Ibid* at para 1 (emphasis added).

maladministration and the abuse of public power, and an explicit acknowledgement that the separation-of-powers concerns can in fact be ‘adequately provided for within the exceptional circumstances test itself’.⁵⁶

The facts of this case can be summarised as follows. In May 2012 the Industrial Development Corporation of South Africa (IDC) issued a request for proposals (RFP) to building contractors for the purposes of prequalifying for a principal building contract to upgrade the IDC head office in Sandton, Johannesburg. The RFP stipulated that late applications would not be evaluated. Seven shortlisted candidates – including Trencon Construction (Pty) Ltd (Trencon, the applicant) and, notwithstanding its late submission, Basil Read (Pty) Ltd (Basil Read, the second respondent) – prequalified pursuant to an assessment of their profiles based on various factors such as technical capability and environmental management.

Already at this prequalification stage, the IDC’s Support Services highlighted to the Procurement Committee that Trencon ‘demonstrated extensive capacity’.⁵⁷ The IDC then commenced the second phase in the tendering process by issuing a formal tender invitation to each of the seven contractors. The tender invitation stated that the site handover date would be 6 September 2012. This second phase was governed by the strict regulatory framework and involved the evaluation of the tenders on the basis of price and broad-based black economic empowerment points in accordance with the 90/10 preference point method. Trencon was awarded 90 points for price given that it submitted the lowest bid price and, pursuant to its empowerment verification certificate, it was also awarded the most points for empowerment. Despite Trencon’s being the clear forerunner at this juncture, clarification was sought by the quantity surveyors on various matters including the question of bid price should the site handover be delayed. Trencon indicated that it would charge an additional monthly escalation amount of 0,6 per cent, whereas Basil Read indicated that its bid would remain fixed regardless. Significantly, despite this price escalation, ‘Trencon’s bid price was still the lowest’⁵⁸ and it was thus recommended (albeit subject to certain conditions) by the quantity surveyors, the IDC’s principal agent for the evaluation of the tender (Snow Consultants Incorporated), the Support Services and in turn the Procurement Committee, which ultimately recommended Trencon for the award of the tender.⁵⁹

Notwithstanding the unanimous endorsement of Trencon’s bid, the IDC’s Executive Committee (Exco, the final decision-maker) declared Trencon’s bid to be non-responsive and awarded the tender to Basil Read.⁶⁰ It reached this decision on the basis that ‘by adding the price escalation, Trencon failed to keep its price fixed for the 120 days of the tender evaluation period’⁶¹ – something the Exco thought was required by the relevant tender documents.

⁵⁶ Ibid at para 94.

⁵⁷ Ibid at para 7.

⁵⁸ Ibid at para 19.

⁵⁹ Ibid at paras 13–20.

⁶⁰ Ibid at para 22.

⁶¹ Ibid.

Trencon challenged this decision in the High Court⁶² and won, essentially on two bases. First, the court held that Basil Read's late proposal should not have been considered insofar as it was a mandatory and material condition of the IDC's RFP that proposals be submitted timeously and that a failure to do so would be a bar to their consideration.⁶³ Secondly, and more importantly, the court held that the final decision was unlawful in that it had been based on a material error of law pursuant to a 'misreading and misunderstanding of ... provisions of the Contract Data ... with cross reference to the JBCC 2000 on price adjustments' in terms of which 'adjustment[s] of the bid price consequent to ... delays in site handover are *not* prohibited'.⁶⁴ This error was conceded by the IDC in argument, and thus it was common cause that Trencon's bid was in fact responsive.⁶⁵ The court therefore set aside the IDC's decision to award the tender to Basil Read.

The High Court then turned to the question of remedy, in particular, whether a substitution order would be just and equitable. It held that a case had indeed been made out to substitute.⁶⁶ In reaching this conclusion, the court seems to have considered the common-law factors cumulatively:

This Court is *qualified* to [take the decision itself]... . [T]he decision was, barring the material error of law, a *foregone conclusion* This tender involves quite a substantial amount of public funds and any further *delay* of the project would cause unjustifiable *prejudice* to Trencon, the IDC and National Treasury... . [I]t will be just and equitable to award the tender to Trencon.⁶⁷

The IDC appealed to the Supreme Court of Appeal (SCA), conceding the material error of law but arguing that 'the court below erred in finding that Basil Read's tender was disqualified because the degree of its submission's lateness was immaterial, caused no prejudice and ought to have been condoned'.⁶⁸ The IDC further contended that the remedy of substitution was inappropriate on the facts.⁶⁹ This latter argument was decisive for the court: having tersely canvassed the principles governing this remedy, Maya JA sweepingly, and without meaningful justification, concluded for a unanimous court that:

[i]t is clear that the court below erred in substituting its own decision... . It overlooked the fact that IDC was *not obliged to award the tender to the lowest bidder or at all*. The award of the tender could not be a *foregone conclusion* in the circumstances. Furthermore, the court

⁶² *Trencon (Pty) Ltd v Industrial Development Corporation of South Africa Ltd and Another* [2013] ZAGPPHC 147.

⁶³ Ibid at para 39.

⁶⁴ Ibid at para 29 (emphasis added).

⁶⁵ Ibid at paras 31–32.

⁶⁶ Ibid at para 53.

⁶⁷ Ibid (emphasis added).

⁶⁸ *Industrial Development Corporation of South Africa Ltd v Trencon Construction (Pty) Ltd and Another* [2014] ZASCA 163, 2014 (4) All SA 561 (SCA) ('Trencon SCA') at para 11.

⁶⁹ Ibid.

does not appear to have balanced the substitution remedy against the *requirements of the separation of powers* and failed to exercise judicial deference.⁷⁰

The court then went on to highlight an ‘additional practical difficulty’ with substitution: ‘Over two years have elapsed since the beginning of the tender process. The information upon which the tenders were evaluated is obviously dated. The order does not accommodate unavoidable supervening circumstances such as price increases that have to be taken into account.’⁷¹ Maya JA thus held that no exceptional circumstances existed to justify substitution and that the appeal therefore succeeded.⁷² In a rigorous, albeit somewhat confusing, unpacking and application of the exceptional circumstances test, the Constitutional Court unanimously reached the contrary conclusion, upholding the appeal by Trencon and reinstating the High Court’s order.⁷³ It is noteworthy that the two highest courts in our land can differ so starkly in their reasoning, and this in turn highlights the value of the appeal process as a ‘check’ on the exercise of judicial power within the separation of powers.

I turn now to explain how the Constitutional Court reaches the conclusion it does and thereby goes some way to clarifying the test for exceptional circumstances where a substitution order is sought. While the Court identifies five particular issues,⁷⁴ for the purposes of this comment I focus on the first three only.

A The Exceptional Circumstances Test

As Lord Steyn put it in the House of Lords decision in *R (on the application of Daly) v Secretary of State for the Home Department*, ‘[i]n law context is everything’.⁷⁵ In unpacking the exceptional circumstances test in *Trencon*, the Court’s analysis is quite evidently informed by both the big-picture context of the need to respect the separation of powers and, more specifically, the context of the wording of s 8(1) of the PAJA.⁷⁶ Thus, against the backdrop of the common-law principles governing the test,⁷⁷ Khampepe J’s starting point is an acknowledgement that ‘[t]he administrative review context of s 8(1) of the PAJA and the wording under subs (1)(c)(ii)(aa) make it perspicuous that substitution remains an extraordinary remedy. Remittal is still almost always the prudent and proper course.’⁷⁸ The Court goes on to elucidate why this is so with reference to the role of the courts within the separation of powers and the concomitant deference required of them

⁷⁰ Ibid at para 18 (emphasis added). Note that Maya JA fails to illustrate how exactly this is ‘clear’. The Court’s analysis on this score is thus extremely thin and the judgment can be criticised on this basis. It should be remembered that judicial officers have a duty to justify their decisions through sound judicial reason-giving. This is an important mechanism by which judges are held accountable within the separation of powers: see M Pieterse ‘Coming to Terms with Judicial Enforcement of Socio-Economic Rights’ (2004) 20 *South African Journal of Human Rights* 383, 391.

⁷¹ *Trencon* SCA (note 68 above) at para 19.

⁷² Ibid at para 21.

⁷³ *Trencon* (note 4 above) at para 101.

⁷⁴ Ibid at para 33.

⁷⁵ *R (on the application of Daly) v Secretary of State for the Home Department* [2001] UKHL 26, [2001] 2 AC 532 (HL).

⁷⁶ See *Trencon* (note 4 above) at paras 35 and 42.

⁷⁷ Ibid at paras 36–41.

⁷⁸ Ibid at para 42 (emphasis added).

in reviewing decisions of administrative agencies, given that 'courts are ordinarily not vested with the skills and expertise required of an administrator'.⁷⁹

Importantly, the Court highlights that '[j]udicial deference, within the doctrine of separation of powers' cuts another way too, insofar as it 'must also be understood in light of the *powers* vested in the courts by the Constitution'.⁸⁰ And when it comes to the courts' remedial powers in constitutional matters, the Constitution is clear: s 172(1) embodies a 'corrective principle' that 'allows correction to the extent of the constitutional inconsistency'.⁸¹ The Court thus notes that '[a] case implicating an order of substitution accordingly requires courts to be mindful of the need for judicial deference *and* their obligations under the Constitution'.⁸² A sensitive balancing is required given a court's dual role of policing constitutional compliance and determining for itself the extent of this policing function in a given case. In the context of awarding substitution orders, Khampepe J proposes the following high-level formulation of the test for exceptional circumstances in an attempt to achieve this balance:

To my mind, given the doctrine of separation of powers, in conducting this enquiry there are *certain factors that should inevitably hold greater weight*. The first is whether a court is in as good a position as the administrator to make the decision. The second is whether the decision of an administrator is a foregone conclusion. These two factors *must* be considered cumulatively. Thereafter, a court should still consider other relevant factors. These may include delay, bias or the incompetence of an administrator. The ultimate consideration is whether a substitution order is just and equitable. This will involve a consideration of fairness to all implicated parties. It is prudent to emphasise that the exceptional circumstances enquiry requires an examination of each matter on a case-by-case basis that accounts for all relevant facts and circumstances.⁸³

I turn now to interpret constructively each of these components of the test as formulated by Khampepe J in the abstract, based on her application of them to the facts.

1 *Separation-of-Powers Factors*

The first thing a court is required to ask in every case is whether it is in 'as good a position', and therefore as well qualified or equipped as the original decision-maker was (and would be again if the matter were to be remitted to him or her). This determination depends on the facts of each case and will thus be informed by various considerations such as (i) *Timing*: Khampepe J states that 'a court ought to evaluate the *stage* at which the administrator's process was situated when the impugned administrative action was taken ... [for] the further along in the process, the greater the likelihood of the administrator having already exercised its specialised knowledge. In these circumstances, a court may very well be in the same position as the administrator to make a decision';⁸⁴ (ii) *Information*: the Court

⁷⁹ Ibid at para 43.

⁸⁰ Ibid at para 45 (emphasis added).

⁸¹ Ibid, citing *AllPay* 2 (note 26 above) at paras 42 and 45.

⁸² *Trencon* (note 4 above) at para 46 (emphasis added).

⁸³ Ibid at para 47 (emphasis added).

⁸⁴ Ibid at para 48 (emphasis added).

seems to emphasise the value of having ‘all relevant information before it’⁸⁵ as an indicator that it may indeed be as well qualified as the administrator and thus able to substitute; and (iii) *The nature of the decision*: at this stage of the overarching enquiry, the Court seems to suggest a theoretical consideration of the type of decision in issue. Thus, in classification- of-functions parlance, a judicial or quasi-judicial-type decision may, provided all relevant information is before the court, place it in as good a position as the administrator to decide the matter.⁸⁶ Then, ‘[o]nce a court has established that it is in as good a position as the administrator, it is competent to enquire into whether the decision of the administrator is a foregone conclusion’.⁸⁷

Establishment of the first factor (‘in as good a position’) is thus a necessary prerequisite for proceeding to the second leg of the test and determining whether the outcome is a ‘foregone conclusion’. On the face of it, this makes sense: how can a court determine whether a particular outcome is inevitable on the facts unless it is in a position, and thus well equipped, to do so? This second mandatory factor in the enquiry necessitates a consideration of whether ‘there is only one proper outcome of the exercise of an administrator’s discretion’,⁸⁸ in other words only one decision that could properly (ie lawfully, reasonably and procedurally fairly) be made, ‘and “it would merely be a waste of time to order the [administrator] to reconsider the matter”’.⁸⁹ This second leg of the test thus essentially requires a determination of whether, given the facts and applicable legislative framework, the administrator has any discretion left to exercise should the matter be remitted, such that a range of possible decisions could properly be made. As Khampepe J puts it, ‘[a] finding that the IDC’s decision is a foregone conclusion depends on whether there was only one proper outcome of the exercise of its discretion and remittal would serve no purpose. In other words, if the matter were to be remitted, the IDC *would not have any discretion left to exercise*’.⁹⁰

The Court goes on to note a practical difficulty that may arise in determining whether an outcome is a foregone conclusion: ‘[I]ndubitably, where the administrator has not adequately applied its unique expertise and experience to the matter, it may be difficult for a court to find that an administrator would have reached a particular decision and that the decision is a foregone conclusion’.⁹¹ However, Khampepe J emphasises that this practical difficulty need not be an insuperable obstacle to this second determination, for ‘where the decision of an administrator is not polycentric and is guided by particular rules or by legislation, it may still be possible for a court to conclude that the decision is a foregone conclusion’.⁹² The second factor thus appears also to turn on a consideration of the

⁸⁵ Ibid. See also at para 58, where the Court places emphasis on the fact that it ‘has the benefit of the record, with all the pertinent information and recommendations’ as an indicator that it was as well-placed as the IDC to make the decision.

⁸⁶ Ibid at para 48, citing *Theron* (note 41 above) at 157B–E.

⁸⁷ Ibid at para 49.

⁸⁸ Ibid.

⁸⁹ Ibid, quoting Hiemstra J in *JCC* (note 42 above) at 76D–H.

⁹⁰ *Trencon* (note 4 above) at para 59 (emphasis added).

⁹¹ Ibid at para 49.

⁹² Ibid.

nature of the decision and the related assessment of whether it raises polycentric and/or policy-laden concerns. Khampepe J herself acknowledges this apparent redundancy in the test: 'The distinction between the considerations in as good a position and foregone conclusion seems opaque', in other words, unclear, 'as they are interrelated and inter-dependent'.⁹³

The Court seems somewhat muddled in attempting to explain this interrelationship and distinctness between the two factors but it seems to me, based on the Court's application of this test to the facts, that at the second stage, rather than taking a broad theoretical approach to the assessment of the nature of the decision, the Court takes a more incisive and practical approach guided by an overarching consideration of whether to defer or not. This second stage of the enquiry appears to turn more on 'the role of policy in the court' rather than on 'a neat list of discrete considerations as is the case at stage one',⁹⁴ and where the nature of the decision is such that policy and/or polycentric concerns necessitate wide-discretionary powers on the part of the decision-maker, a court would be ill-placed and unwise to determine a decision to be a forgone conclusion.

Regarding these first two factors, which go to the institutional competence of the court within the separation of powers, Khampepe J is emphatic in stating that they 'should inevitably hold greater weight'⁹⁵ in the overall enquiry. However, an unfortunate shortcoming of the judgment is that it is not entirely clear on a literal interpretation whether this means that they must simply weigh more in the balance when considering the other factors, or whether they are strict requirements that must be met in each case for substitution to be ordered: in other words, that establishing one of the other factors (such as bias) on its own will no longer be sufficient to justify substitution. It seems to me that the latter position is intended to be the case. This conclusion is supported by a purposive understanding of the test as set out by the Court against the backdrop of Khampepe J's analysis of the need for '[j]udicial deference, within the doctrine of separation of powers'.⁹⁶ This assessment is, in turn, bolstered by the judge's emphatic insistence that 'the separation of powers is adequately provided for within the exceptional circumstances test itself',⁹⁷ and it accords with the foundations laid by Plasket J in *Intertrade*, where he indicated that the institutional expertise and competence of the court are 'necessary prerequisites'⁹⁸ in the overarching exceptional circumstances enquiry.

The recent SCA judgment in *Westinghouse* also appears to endorse this interpretation of the *Trencon* formulation: 'The [Constitutional Court] said ... that the first enquiry is whether it is in as good a position to make the decision as the administrator was. Second, it *must* determine whether a substituted award is a foregone conclusion.'⁹⁹ On this interpretation, *Trencon* presents a noteworthy

⁹³ Ibid at para 50.

⁹⁴ Professor Hugh Corder's comments to me in a discussion about this case.

⁹⁵ *Trencon* (note 4 above) at para 47.

⁹⁶ Ibid at para 45.

⁹⁷ Ibid at para 94.

⁹⁸ *Intertrade* (note 5 above) at para 43.

⁹⁹ *Westinghouse Electric Belgium SA v Eskom Holdings (Soc) Ltd and Another* [2015] ZASCA 208, 2016 (3) SA 1 (SCA), 2016 (1) All SA 483 (SCA) at para 74 (emphasis added).

development in our law: notwithstanding the existence of other factors such as bias, a court must engage in a cumulative consideration of the separation-of-powers factors and find itself well equipped and thus competent to decide a matter before an order of substitution can be made. While this certainly provides greater structure to the enquiry, it should be noted that this kind of cumulative assessment is not a one-size-fits-all solution. Thus, in seeking to accommodate the separation-of-powers concerns within the test, Khampepe J has arguably made it harder for litigants to meet the case for substitution in certain instances, namely where the separation-of-powers requirements cannot be met but the facts, which evidence for example glaring incompetence or bias, nonetheless cry out for substitution.

2 *Other Factors such as Bias, Gross Incompetence, Delay and Good Faith*

Although a consideration of the separation-of-powers factors appears to be mandatory and their presence would seem to be a *sine qua non* for substitution, this is not the end of the enquiry: '[a] court *must* consider other relevant factors'¹⁰⁰ which may carry some weight in the balance. These essentially equitable considerations include the established grounds of bias and incompetence¹⁰¹ – although in relation to the latter, Khampepe J seems to set a higher standard than that required at common law by the addition of the qualifier 'gross' – as well as delay. In her application of the test under the heading 'other considerations', she also adds a new consideration, that of good faith on the part of the decision-maker.¹⁰² This may weigh in favour of remittal, and conversely it would seem that evidence of bad faith may strengthen the case for substitution.

On the subject of delay, the Court makes some important remarks. First, it is highlighted that given the contextual nature of the exceptional circumstances test, '[d]elay can cut both ways'¹⁰³ as a factor. Sometimes the facts may show that the result of delay is, for example, 'a drastic change of circumstances [such that] a party is no longer in a position to meet the obligations arising from an order of substitution'; or perhaps they indicate that 'the needs of the administrator have fundamentally changed'.¹⁰⁴ In such scenarios, the practical realities occasioned as a result of the delay would render a determination as to 'foregone conclusion' highly unlikely and so support remittal over substitution. Conversely, the facts may be indicative of delay warranting substitution, for example 'where a party is prepared to perform in terms of that order and has already suffered prejudice by reason of delay'.¹⁰⁵ In such instances, the Court emphasises that delay occasioned by remittal 'may very well result in further prejudice to that party', and importantly, it 'may also negatively impact the public purse'.¹⁰⁶ This latter remark is a subtle example of the Court's being guided in its unpacking of

¹⁰⁰ *Trencon* (note 4 above) at para 51.

¹⁰¹ *Ibid* at paras 53–54, where Khampepe J couples a consideration of these factors with the notion of fairness.

¹⁰² *Ibid* at para 78.

¹⁰³ *Ibid* at para 51.

¹⁰⁴ *Ibid*.

¹⁰⁵ *Ibid*.

¹⁰⁶ *Ibid*.

the test by the constitutional prescripts of public procurement (fairness, equity, transparency, competitiveness and cost-effectiveness),¹⁰⁷ and indeed by the public interest more generally.

The second important point the Court makes about the delay factor is that 'delay occasioned by the litigation process should not easily cloud a court's decision in reaching a just and equitable remedy'.¹⁰⁸ The reasons for this are both principled and pragmatic. First, as a matter of principle, Khampepe J notes that 'an appeal should ordinarily be decided on the facts that existed when the original decision was made'.¹⁰⁹ In other words, '[d]elay must be understood in the context of the facts that would have been laid in the court of first instance as that is the court that would have been tasked with deciding whether a substitution order constitutes a just and equitable remedy in the circumstances'.¹¹⁰ The related pragmatic reason for this approach is that delay is an inevitable outcome of the appeal process, and thus 'assessing delay with particular reference to the time between the original decision and when the appeal is heard could encourage parties to appeal cases ... with the hope that the time that has lapsed in the litigation process would be a basis for *not* granting a substitution order'.¹¹¹ The Court seems loath to adopt such an approach on the implicit basis that it would handicap the courts in exercising their wide discretionary powers to grant just and equitable relief, including, where appropriate, orders of substitution. Thus Khampepe J emphasises that where a litigant wishes to raise concerns flowing from the delay occasioned by the appeal process, any such 'new evidence ... must be adduced and admitted in accordance with legal principles applicable to the introduction of new evidence on appeal'.¹¹²

3 *The Role of Fairness in the Enquiry*

The final step in the exceptional circumstances test involves an assessment of which way fairness tips the substitution scales. Thus, the Court confirms that '*ultimately*, the appropriateness of a substitution order must depend on the consideration of fairness to the implicated parties'.¹¹³ In other words, what Khampepe J seems to suggest is that the specific factors delineated above cannot be considered in the abstract, given the context-sensitive nature of the remedy and the overarching dictates of justice and equity. In this way the Court attempts to strike the requisite balance between ensuring a degree of both certainty and flexibility in the exceptional circumstances test, for neither extreme (hard-and-fast rules versus utter vagueness) is a good thing. Khampepe J then provides an example of how fairness might tip the scales: where an administrator 'is found to have been biased or grossly incompetent, it may be unfair to ask a party to resubmit itself to the administrator's jurisdiction... . However, having regard to the notion of fairness, a court may still substitute even where there is no instance

¹⁰⁷ These procurement principles are encapsulated in s 217 of the Constitution (emphasis added).

¹⁰⁸ *Trencon* (note 4 above) at para 52.

¹⁰⁹ *Ibid.*

¹¹⁰ *Ibid.*

¹¹¹ *Ibid* at para 53 (emphasis added).

¹¹² *Ibid.*

¹¹³ *Ibid* (emphasis added).

of bias or incompetence'.¹¹⁴ It is significant to note that this position does not appear to hold in relation to the separation-of-powers factors. In other words, it seems that it will never be just and equitable (or simply, fair) to substitute in the absence of their establishment,¹¹⁵ a development that may prove to be troublesome for litigants in certain instances. Given the significance of this development, it is a pity that Khampepe J was not clearer and more explicit regarding the interplay between the overarching consideration of fairness and the other factors in the enquiry.

Although the Court seemingly places particular emphasis on the role of fairness, I do not think this is especially novel. As Hoexter has remarked, '[f]airness to both sides has always been and will almost certainly remain an important consideration'.¹¹⁶ Its role in the enquiry is necessitated by 'the flexibility embedded in the notion of what is just and equitable'.¹¹⁷

In conclusion, having sought to clarify the test, Khampepe J remarks that the version of it presented is 'consonant with the Constitution while at the same time giving proper deference and consideration to an administrator'.¹¹⁸ For what it is worth, my view is that despite the lack of clarity in parts, the Court indeed manages to strike this delicate balance. I turn now to elucidate how the Court applies the test to the facts of the case.

B Application of the Test to the Facts in *Trencon*

The key issue before the *Trencon* Court was whether the circumstances of the case were exceptional and thus warranted a substitution order.¹¹⁹ Through a rigorous application of the test to the unique factual matrix of the case, the Court indeed finds this to be so.

1 In as Good a Position

First, the Court establishes that it 'is in *as good a position* as the IDC to award the tender to *Trencon*'.¹²⁰ This conclusion is reached on the following basis. In terms of the timing consideration, Khampepe J notes that '[t]he material error of law occurred when the procurement process was in the stages of finalisation' and all the relevant bodies 'had considered the bids and undertaken all the technical components of the process'.¹²¹ All that remained was for Exco to approve the recommendation of the Procurement Committee and, significantly, '[t]he IDC itself stated that Exco had fully considered *Trencon*'s bid'.¹²² Then, regarding the consideration of requisite information, Khampepe J notes that the Court 'has the benefit of the record, with all the pertinent information and recommendations

¹¹⁴ Ibid at para 54.

¹¹⁵ See *Trencon* (note 4 above) at para 47.

¹¹⁶ Hoexter (note 37 above) at 553.

¹¹⁷ *Trencon* (note 4 above) at para 55.

¹¹⁸ Ibid.

¹¹⁹ See ibid at para 56.

¹²⁰ Ibid at para 57 (emphasis added).

¹²¹ Ibid.

¹²² Ibid at para 58.

that would have been before Exco'.¹²³ Finally – and here is where we see the interrelationship and arguable overlap between the two separation-of-powers factors play out practically – in considering the nature of decision and whether polycentricity may still come into the equation, Khampepe J notes that 'the IDC does not explain how its administrative expertise could come into play at this point or on what basis it could decide differently'.¹²⁴ Having established that it is in as good a position, and thus as well qualified as the IDC to decide, the Court moves on to a determination of whether the award of the tender to Trencon is a foregone conclusion.

2 *Foregone Conclusion*

This second mandatory requirement of the test is also found to be met, essentially because on remittal, the IDC would have no discretion left to exercise.¹²⁵ This conclusion is informed by the accepted facts that Trencon was the clear forerunner at both stages of the procurement process and that all relevant internal and external expert bodies had recommended that it be awarded the tender.¹²⁶ Furthermore, Khampepe J highlights that 'but for an error of law regarding Trencon's price escalation for the delayed site handover, Trencon's bid would not have been declared non-responsive'.¹²⁷ There could therefore be only one proper outcome of the exercise of the IDC's discretion. It is on this score that the Court differs starkly with the SCA which, 'despite finding that the IDC could not have lawfully awarded the tender to any other bidder',¹²⁸ found that remittal was the proper course 'on the basis that the IDC still had a discretion not to award the tender to the highest points earner or not to proceed with the tender at all'.¹²⁹ The IDC persisted with this 'discretion' argument before the Constitutional Court, advancing three particular contentions – each of which is strongly refuted in a manner suggesting abhorrence to unconstrained and/or unlawful exercises of discretionary power.

First, the IDC claimed it had a discretion not to award the tender to the highest points-earner.¹³⁰ The Court found this argument to be misplaced. This was essentially because this broad discretion is curtailed by the relevant provisions of the regulatory framework read with the Standard Conditions of Tender. In particular, s 2(1)(f) of the Preferential Procurement Policy Framework Act 5 of 2000, read with clause F.3.11.3(d) of the Standard Conditions of Tender, mandate the IDC to award the tender to the highest points-earner save where there exist 'objective criteria' or 'justifiable reasons' for not doing so.¹³¹ The various concerns pertaining to Trencon's bid price did not qualify as such objective criteria for, their existence notwithstanding, Trencon's bid was still lower than all the other bids

¹²³ Ibid.

¹²⁴ Ibid.

¹²⁵ Ibid at para 59.

¹²⁶ Ibid.

¹²⁷ Ibid.

¹²⁸ Ibid at para 60.

¹²⁹ Ibid.

¹³⁰ Ibid at para 62.

¹³¹ Ibid.

and any disputes regarding these pricing matters could be dealt with contractually under the JBCC Agreement.¹³² This conclusion was fortified by Exco's own admission that the only reason given by the IDC for refusing Trencon the award was the material error of law.¹³³ Given the peremptory language of the regulatory framework, in the absence of objective criteria the IDC could not possibly be justified in not awarding the tender to the highest points-earner, Trencon.¹³⁴

Secondly, the IDC claimed that it did not have a proper opportunity to evaluate all the bids as its evaluation was tainted by the error of law.¹³⁵ The Court does not hesitate to dismiss this 'disingenuous'¹³⁶ argument that was so patently undercut by contradictions in the IDC's own version of events. In this regard, Khampepe J points out that the IDC itself 'proclaimed that "the decision of Exco" was arrived at [by] taking into account the totality of facts before Exco' which, by its own admission, "applied its mind to issues which were relevant in relation to its decision [to award the tender]", and furthermore, the IDC definitively asserted in its affidavit that there was 'no evidence that Trencon's tender was "not properly evaluated"'.¹³⁷ The remarkable attempt to claim that Exco had not had a proper opportunity to apply its mind to the bid therefore had to fail.

Lastly, in attempting to show that the award to Trencon was not a foregone conclusion, the IDC asserted that it had a discretion not to award the tender at all.¹³⁸ Again, it was forced to concede on appeal that this discretion was not uncurtailed. In particular, reg 8(4) of the Procurement Framework Regulations (2011) empowers an organ of state to cancel a tender prior to its award only in three circumscribed instances: (i) where due to changed circumstances there is no longer a need for the services, works or goods in question; or (ii) where there are no longer sufficient funds to cover the envisaged total expenditure; or (iii) where no acceptable tenders are received.¹³⁹ The evidence patently showed that none of these grounds was present. The IDC clearly intended to continue with the tender and had sufficient funds to do so – as Khampepe J so poignantly points out, '[t]he fact that the IDC ultimately awarded the tender to Basil Read provides sufficient credence for this'¹⁴⁰ – and there was simply 'no basis for the IDC to argue that no acceptable tenders were received'.¹⁴¹ Furthermore, the wording of the regulations is clear: cancellation can happen only *prior* to the actual award of a tender. Thus the IDC, having already awarded the tender to Basil Read, could not belatedly attempt to exercise this power to cancel.¹⁴² In this respect, the Court finds that the SCA erred 'in conceiving that the contractual power not to award the tender at all could in these circumstances have been lawfully exercised'.¹⁴³ Finally, in

¹³² See *ibid* at paras 63–64.

¹³³ *Ibid* at para 65.

¹³⁴ See *ibid*.

¹³⁵ See *ibid* at para 66.

¹³⁶ *Ibid*.

¹³⁷ *Ibid*.

¹³⁸ See *ibid* at para 68.

¹³⁹ See *ibid*.

¹⁴⁰ *Ibid* at para 69.

¹⁴¹ *Ibid*.

¹⁴² See *ibid* at para 71.

¹⁴³ *Ibid*.

refuting this last leg of the IDC's discretion argument, Khampepe J makes a noteworthy observation clearly underpinned by a need to protect the courts' wide remedial powers and, in turn, their policing function within the separation of powers:

If, when arguing that remittal is the proper remedy, an organ of state is able to raise the fact that it has this discretion *without more*, a court would virtually never have the power to grant a remedy of substitution. The organ of state would always be able to argue that it still had a discretion not to award the tender, *thereby constraining the power of the courts to grant just and equitable remedies*. It is a fundamental principle of the rule of law that organs of state, like the IDC, can only exercise power that has been conferred onto them. They cannot, on their own volition, confer power unto themselves that was never there.¹⁴⁴

In light of all of the foregoing reasons, the Court concludes that 'the award of the tender to Trencon is a foregone conclusion'.¹⁴⁵ Having established the separation-of-powers factors, the Court goes on to consider other relevant factors that might weigh in the balance in determining whether substitution would be just and equitable.

3 *Delay and Supervening Circumstances*

As a result of the litigation, two years had elapsed since the commencement of the tender process. The SCA found substitution to be inappropriate given that prices might have increased during this period.¹⁴⁶ Rather robustly, Khampepe J holds that '[o]n this point, the Supreme Court of Appeal erred'.¹⁴⁷ The principled reasons propounded for this finding are twofold. First, the SCA should have determined the matter on the basis of the facts that were before the court of first instance instead of 'on the basis of the delay incurred as a result of the appeal itself';¹⁴⁸ and insofar as this delay did indeed lead to supervening circumstances, no new evidence had been adduced to show as much.¹⁴⁹ A related point, which flows from the fact that the delay factor cuts two ways, is that account must be taken of any adverse impact on the public purse that would be caused by 'the further delay occasioned by remittal'.¹⁵⁰ This consideration flows from a principle formulated thus by Khampepe J: 'Procurement disputes, especially those involving organs of state, must be resolved expediently'.¹⁵¹

The second reason the Court finds the SCA to have erred in its application of the delay factor is that it 'did not value the distinction between public and private law'¹⁵² insofar as the decision to award a tender is a public-law matter strictly regulated by the legislative framework and, in comparison, matters such as contract price adjustment which are subject to negotiation post-award, 'ought to

¹⁴⁴ Ibid at para 70 (emphasis added).

¹⁴⁵ Ibid at para 71.

¹⁴⁶ See ibid at para 72.

¹⁴⁷ Ibid.

¹⁴⁸ Ibid at para 73.

¹⁴⁹ See ibid.

¹⁵⁰ Ibid at para 74.

¹⁵¹ Ibid.

¹⁵² Ibid at para 75.

fall *squarely* within the domain of private law'.¹⁵³ Given that the JBCC Agreement provides for such price adjustments, this option would be 'viable to account for the delay in these circumstances'¹⁵⁴ and so, especially given that both parties conceded that negotiations post-award are subject to private law, '[t]he Supreme Court of Appeal erred in revoking the High Court's order of substitution on this ground'.¹⁵⁵

A further argument raised by the IDC relating to the matter of delay was 'that substitution is never appropriate where the tender validity period has expired'.¹⁵⁶ Two cases were referred to in support of this contention.¹⁵⁷ However, in both of them the relevant organ of state had not yet awarded the tender, and in the present instance the IDC had clearly done so. Consequently, remarks Khampepe J, 'a substitution order here would not require the tender validity period to be extended because this period is held in abeyance pending the finalisation of the matter'.¹⁵⁸ The judge proceeds to provide a principled and pragmatic, foundation for this conclusion informed by a purposive understanding of the objects of the PAJA:

Once an award has been challenged, the litigation process will inevitably run longer than the 120-day tender validity period. If [this] period, in itself, were to be treated as a bar to an order of substitution, there may be no incentive for an aggrieved party like Trencon to lodge review proceedings. This is because its desired remedy – that of substitution – would not be available to it. This approach would not accord with the objectives of PAJA as the tender validity period would, in most instances, be deemed to have expired. Courts would almost always be deprived of their powers of substitution.¹⁵⁹

Again, this dictum serves as a captivating example of the Court affirming the constitutional imperative that it not be hampered in exercising its powers to award appropriate relief. Khampepe J's approach here seems motivated by the need to ensure aggrieved bidders are not disincentivised from challenging apparently unlawful tender awards, and the related need to protect the judicial space required to scrutinise public procurement through the lens of the constitutional and regulatory prescripts.

¹⁵³ Ibid (emphasis added). This is quite a remarkable statement and seems to me to be an oversimplification of matters: surely one cannot insulate post-award negotiation from the public-law constraints of, for example, the constitutional procurement principles? In light of these principles, Bolton has remarked that '[post-award] [n]egotiations are only allowed if the preferred tenderer will remain the most favoured tenderer in accordance with the tender criteria, and the contract will not be significantly different from the contract initially advertised. More or less similar rules also apply to the variation of a contract after its conclusion – changes made may not result in a contract that is materially different from the contract initially advertised.' This is because of public-law considerations such as fairness: P Bolton 'The Scope for Negotiating and/or Varying the Terms of Government Contracts Awarded by way of a Tender Process' (2006) 12 *Stellenbosch Law Review* 266, 287.

¹⁵⁴ *Trencon* (note 4 above) at para 76.

¹⁵⁵ Ibid at para 77.

¹⁵⁶ Ibid at para 79.

¹⁵⁷ *Telkom SA Ltd v Merid Training (Pty) Ltd and Others; Bibati Solutions (Pty) Ltd v Telkom SA Ltd and Others* [2011] ZAGPHC 1 and *Joubert Galpin Searle Inc and Others v Road Accident Fund and Others* [2014] ZAECPEHC 19, 2014 (4) SA 148 (ECP).

¹⁵⁸ *Trencon* (note 4 above) at para 80.

¹⁵⁹ Ibid at para 81.

4 *Other Considerations and the Role of Fairness*

Finally, the Court takes note of the fact that 'the IDC acted in good faith when it was moved by a material error of law', which 'should be a strong consideration for the Court when considering whether to grant a substitution order'.¹⁶⁰ Ultimately, however, given the findings in relation to the other relevant factors and 'viewed through the lens of fairness to both parties', Khampepe J holds that 'it would be unfair for this Court to remit the matter to the IDC'.¹⁶¹ She concludes that '[t]he unique circumstances of this case present a good example, in administrative law, of an instance where the Court is not usurping the functions of the administrative body by making a substitution order'.¹⁶² The Court thus finds that such an order was indeed just and equitable in the circumstances.¹⁶³

V CONCLUSION

The Court in *Trencon* had to confront head-on, for the very first time, the appropriate calibration of the exceptional circumstances test where an order of substitution is sought. This was done against the backdrop of the separation of powers and the related degree of deference required of courts given the review / appeal divide. Substitutive decisions epitomise the disintegration of this divide, and thus it is only exceptionally that judicial review will give rise to such decisions. As the oft-quoted 'Mureinikism' goes, our Constitution has necessitated a move from a 'culture of authority' to a 'culture of justification' pursuant to which 'every exercise of power is expected to be justified'.¹⁶⁴ This includes the potentially expansive power of judicial review. Courts must thus be justified in substituting their decisions for those of administrators skilled in managing the complexities of polycentric matters.

The Court in *Trencon* seeks to clarify exactly when the courts will indeed be justified in doing so. There is immense value in the highest court addressing this issue, and with such rigour. For this reason, the judgment is noteworthy for the principles it espouses and, more than this, it will no doubt have important practical spillover effects. In providing, for the most part, a greater degree of clarity and detail in the formulation of the test, the Court has empowered potential litigants better to determine whether to challenge apparently unlawful tender awards given that the prospect of substitution, and the tangible relief it presents, may be a strong incentive for doing so. In this respect the Court pays homage to the constitutional procurement principles and seems to have at heart the public interest in ensuring their fulfilment. The judgment is thus illustrative of the significant role that the courts play as a 'check' within the separation of powers, as guardians of the Constitution, and as watchdogs to deter and address maladministration, which, as was so poignantly shown by the IDC's Exco, remains an unfortunately commonplace occurrence in South Africa. At the same

¹⁶⁰ Ibid at para 78.

¹⁶¹ Ibid.

¹⁶² Ibid at para 98.

¹⁶³ See *ibid*.

¹⁶⁴ E Mureinik 'A Bridge to Where? Introducing the Interim Bill of Rights' (1994) 10 *South African Journal of Human Rights* 31, 32 (emphasis added).

time, in putting centre-stage those factors that go to the institutional expertise and competence of a court in the formulation of the test, Khampepe J ensures respect for the degree of separation required by the doctrine of separation of powers.

For these reasons, in constructively interpreting the principles set out in this judgment I have sought to illustrate that it goes some way to achieving that ever delicate balance between ensuring neither complete judicial reserve nor excessive boldness; neither inflexible rules, nor utterly unguided discretion. When it comes to the question of remedy, this much is demanded by the constitutional dictates of justice and equity.

Clarifying the Exceptional Circumstances Test in *Trencon*: An Opportunity Missed

Raisa Cachalia*

I INTRODUCTION

In *Trencon* the Constitutional Court undertook to clarify the test for exceptional circumstances in s 8(1)(c)(ii)(aa) of the Promotion of Administrative Justice Act 3 of 2000 (PAJA) for determining whether a court should substitute an administrator's decision with its own.¹ In relevant part, s 8(1)(c) of the PAJA reads:

The court or tribunal, in proceedings for judicial review ... may grant any order that is *just and equitable*, including orders—

...

- (c) setting aside the administrative action and—
 - (i) remitting the matter for reconsideration by the administrator, with or without directions; or
 - (ii) in *exceptional cases*—
 - (aa) *substituting or varying the administrative action* or correcting a defect resulting from the administrative action.²

While Khampepe J acknowledged that the wording of s 8(1) read with s 8(1)(c)(ii)(aa) makes substitution an 'extraordinary remedy' and that remittal is, for this reason, 'still almost always the prudent and proper course',³ her complicated (and even contradictory) formulation of the test and her proclivity for fairness dilute the exceptional nature of the remedy and do little to clarify when substitution is warranted. In the context of failing public procurement, it is perhaps understandable that courts would more readily come to the assistance of parties aggrieved by defective administrative processes. Khampepe J suggested as much in *Trencon* when she alluded to the 'vital role' that tendering plays in the delivery of goods and services to the South African public; the '[l]arge sums of public money' involved; and the immense power the government wields in making

* BA LLB LLM (Witwatersrand); Researcher, SAIFAC (South African Institute for Advanced Constitutional, Public, Human Rights and International Law), University of Johannesburg; attorney of the High Court of South Africa and legal consultant at Caveat Legal (Pty) Ltd. I am deeply indebted to Leo Boonzaier for his meticulous comments and advice, and to Lauren Kohn for some healthy debate about the case.

¹ *Trencon Construction (Pty) Ltd v Industrial Development Corporation of South Africa Ltd and Another* [2015] ZACC 22, 2015 (5) SA 245 (CC), 2015 (10) BCLR 1199 (CC) ('*Trencon*') at para 32.

² Emphasis added.

³ *Trencon* (note 1 above) at para 42.

tender awards.⁴ But as Plasket J cautioned in *Intertrade*, and s 8(1) of the PAJA makes clear, overzealous judicial intervention is not the panacea for all cases of administrative breakdown:

[C]ourts, when considering the validity of administrative action, must be wary of intruding, even with the best of motives, without justification into the terrain that is reserved for the administrative branch of government. These restraints on the powers of the courts are universal in democratic societies such as ours and necessarily mean that there are limits on the powers of the courts to repair damage that has been caused by a breakdown in the administrative process.⁵

This warning explains why we require a guided approach to determining whether substitution is justified. In her note on the same case, Lauren Kohn concludes that the *Trencon* judgment goes ‘some way’ in achieving the delicate balance between certainty and flexibility.⁶ But it is on this point that we diverge. While it is true that legal rules are no longer the be-all and end-all of our legal system, and courts are plainly required to balance rulemaking with the more elastic, normative demands of the Constitution,⁷ my argument is that in its drive for flexibility the Court has unjustifiably sacrificed certainty and the value of providing guidance to lower courts and future litigants on how to exercise the discretion in question. In the words of Wallis: ‘High-flown rhetoric and sonorous phraseology are no substitute for principled analysis and reasoning and clarity of expression.’⁸

II THE TRADITIONAL APPROACH TO SUBSTITUTION

It is a generally accepted principle of our common law that a court will be reluctant to assume decision-making power for itself where the discretion has been entrusted to another functionary.⁹ This flows from the separation of powers¹⁰ and, most crucially, from the appreciation of the proper purpose of judicial review, which is ‘to scrutinise the legality of administrative action, not to secure a decision by a judge in place of an administrator’.¹¹ For this reason, and barring special reasons (or ‘exceptional cases’ under the PAJA), the courts have

⁴ Ibid at para 1.

⁵ *Intertrade Two (Pty) Ltd v MEC for Roads and Public Works, Eastern Cape, and Another* [2007] ZAECHC 149, 2007 (6) SA 442 (Ck), [2008] 1 All SA 142 (Ck) (‘*Intertrade*’) at para 46.

⁶ L Kohn ‘The Test for “Exceptional Circumstances” where an Order for Substitution is Sought: An Analysis of the Constitutional Court Judgment in *Trencon* against the Background of the Separation of Powers’ in this volume at 91.

⁷ Constitution of the Republic of South Africa, 1996.

⁸ M Wallis ‘Commercial Certainty and Constitutionalism: Are they Compatible?’ 22, available at <http://scholar.ufs.ac.za:8080/xmlui/bitstream/handle/11660/1789/WallisM.pdf?sequence=1>.

⁹ See, eg, *Premier, Mpumalanga and Another v Executive Committee, Association of State-Aided Schools, Eastern Transvaal* [1998] ZACC 20, 1999 (2) SA 91 (CC), 1999 (2) BCLR 151 (CC) (‘*Premier, Mpumalanga*’) at para 50 and *Johannesburg City Council v Administrator, Transvaal and Another* 1969 (2) SA 72 (T) (‘*JCC*’) at 76E.

¹⁰ See, eg, *Premier, Mpumalanga* (note 9 above) at para 51; *Intertrade* (note 5 above) at para 45; *International Trade Administration Commission v SCAW South Africa (Pty) Ltd* [2010] ZACC 6, 2012 (4) SA 618 (CC), 2010 (5) BCLR 457 (CC) (‘*SCAW*’) at para 95 and *Doctors for Life International v Speaker of the National Assembly and Others* [2006] ZACC 11, 2006 (6) SA 416 (CC), 2006 (12) BCLR 1399 (CC) (‘*Doctors for Life*’) at para 37.

¹¹ L Baxter *Administrative Law* (1984) 481.

reiterated that the proper course is ‘almost always’¹² to remit the matter to the decision-maker for reconsideration.¹³

Deciding when to substitute has persistently vexed the courts. This power has, however, been recognised in circumstances where: (i) the end result is a foregone conclusion such that remittal would be a ‘mere formality’ or ‘waste of time’ given the inevitability of the outcome;¹⁴ (ii) there is a delay causing unjustifiable prejudice to the affected party;¹⁵ (iii) bias or incompetence on the part of the administrator is established such that ‘it would be unfair to require the applicant to submit to the same jurisdiction again’;¹⁶ or (iv) where the court finds itself in ‘as good a position’ as the administrator to take the decision itself.¹⁷ What has not been clear or consistent is how factors (i)–(iv) interact: is any factor sufficient on its own to justify substitution or are the factors sufficient only in combination with other factors? Are certain factors necessary preconditions for substitution? Or are they all simply weighed together in deciding whether there is an ‘exceptional case’ justifying substitution?

A The Relationship between the Factors

In *Johannesburg City Council v Administrator, Transvaal (JCC)*, the *locus classicus* on substitution, the Johannesburg City Council sought approval from the provincial administrator to erect a power station on a farm in the area.¹⁸ Prior to granting an approval the administrator was obliged to call upon the Electricity Supply Commission (Escom) to furnish a report on the proposal. During this process, Escom offered to supply the required electricity itself, leaving the administrator to decide which of the options would be most advantageous. The council’s application was twice refused in favour of Escom and both times the refusal was reviewed and set aside. However, on the second occasion, the Appellate Division was tasked with deciding whether to take the decision itself or refer it back to the administrator for reconsideration. Notwithstanding that the administrator’s errors had already resulted in a three-year delay – and that he had been twice mistaken – the court was not satisfied that the council’s case for self-generation of electricity was so ‘overwhelmingly strong that no reasonable man could decide otherwise’.¹⁹ In essence, the court found that mere delay (factor (ii)) would be insufficient on its own to warrant substitution and that the result must, in addition, be a foregone conclusion (factor (i)).

There are, however, cases where the courts have suggested that delay could be an independent basis for substitution. An example is *Ruyobezza*, a case about a refugee’s request for a certificate of clearance from the Standing Committee on

¹² *Gauteng Gambling Board v Silverstar Development Ltd and Others* [2005] ZASCA 19, 2005 (4) SA 67 (SCA) (‘*Silverstar*’) at para 1.

¹³ See *JCC* (note 9 above) at 76E.

¹⁴ *Ibid* at 76E–G; Baxter (note 11 above) at 682.

¹⁵ *JCC* (note 9 above) at 76F; *M v Minister of Home Affairs and Others* [2014] ZAGPPHC 649 at paras 166 and 175–176.

¹⁶ See, eg, *JCC* (note 9 above) at 76F–G.

¹⁷ *Silverstar* (note 12 above) at paras 28 and 39; Baxter (note 11 above) at 681–684.

¹⁸ Note 9 above.

¹⁹ *Ibid* at 77E.

Refugee Affairs, which is required for the applicant's subsequent application for an immigration permit.²⁰ Although Thring J remarked that evidence had been presented pertaining to the 'the high level of violence still prevalent in Burundi',²¹ the court's resolution to substitute its decision for that of the committee seems to have been based entirely on the prejudice and unfairness of the present and future delays (factor (ii)).²² In response to *Ruyobezu*, however, in *Radjabu Binns-Ward J* concluded that:

Issues such as the prejudice occasioned by delay ... cannot justify the granting of asylum in circumstances in which it is not *sufficiently clear* that an applicant qualifies for refugee status in terms of s 3 of the Refugees Act.²³

For this, the judge relied on the reasoning of Murphy J in *Tantoush*.²⁴ While in that instance Murphy J accepted that there had been both bias and prejudicial delay, he said that the most important reason why substitution was warranted was that the evidence before the court demonstrated that the applicant qualified for refugee status.²⁵ It was unclear whether Binns-Ward J and Murphy J respectively regarded the fact that the applicants qualified for refugee status as indicating that remittal would serve no purpose because the outcome was a foregone conclusion (factor (i)), or whether the courts felt that they were in as good a position as the relevant refugee body to take the decision (factor (iv)). Either way, both judgments suggest that something more than delay is needed to justify substitution. What these cases show is that a mere delay – without the outcome being a foregone conclusion or the court being in as good a position – will not be a sufficient basis for substitution.

In *Competition Commission*, following an unsatisfactory decision regarding an application for an exemption under the Competition Act 89 of 1998, the Supreme Court of Appeal (SCA) refused to accept the complainant's argument that the court a quo 'was in as good a position as the Commission to grant or refuse exemption and that, *for this reason alone*, the matter was rightly not remitted'.²⁶ The court cited with approval Baxter's view that '[t]he mere fact that a court considers itself as qualified to take the decision as the administrator *does not of itself*

²⁰ *Ruyobezu and Another v Minister of Home Affairs and Others* 2003 (5) SA 51 (C), 2003 (8) BCLR 920 (C) ('*Ruyobezu*'). See also *Airoadexpress (Pty) Ltd v Chairman of the Local Road Transportation Board, Durban and Others* [1986] ZASCA 6, 1986 (2) SA 663 (A) ('*Airoadexpress*') at 680F–G.

²¹ *Ruyobezu* (note 20 above) at 54F.

²² *Ibid* at 65C–H. See also *Reizis NO v MEC for the Department Sport, Arts, Culture and Recreation and Others* [2013] ZAFSHC 20 ('*Reizis*') at para 38, where the court, relying on *Ruyobezu* (note 20 above), seemed to regard delay as an independent factor. See further *Reynolds Brothers Ltd v Chairman, Local Road Transportation Board, Johannesburg and Another* 1985 (2) SA 790 (A) at 805F–H concerning a refusal to grant certain private road-carrier permits. Miller JA, in directing that the desired permits be issued, said that it was of 'considerable importance ... in the organisation and running of [the appellant's] business' that there be 'clarity and assurance as to the means of conveyance of its products'.

²³ *Radjabu v Chairperson of the Standing Committee for Refugee Affairs and Others* [2014] ZAWCHC 134, 2015 (1) All SA 100 (WCC) ('*Radjabu*') at para 34 (emphasis added).

²⁴ *Tantoush v Refugee Appeal Board and Others* [2007] ZAGPHC 191, 2008 (1) SA 232 (T) ('*Tantoush*').

²⁵ *Ibid* at para 128.

²⁶ *Commissioner, Competition Commission v General Council of the Bar of South Africa and Others* [2002] ZASCA 101, 2002 (4) All SA 145 (SCA) ('*Competition Commission*') at para 15 (emphasis added).

justify usurping the administrator's power'.²⁷ Thus, on the court's reasoning, factor (iv) (in as good a position) would not, without more, justify substitution. However, by remarking that '[a] reasonable apprehension of bias on the part of an administrator is, of course, an established ground for refusing a remittal',²⁸ the SCA seemed to be saying that bias (factor (iii)), unlike factor (iv), could be an independent justification for substitution. This appears also to have been the view taken by the courts in *JCC*²⁹ and *Erf 167 Orchards*.³⁰

However, in *Harerimana*, notwithstanding the court's suggestion that bias on its own would justify substitution, Davis J went on to consider what decision the court could make practically given the information before it.³¹ This appears to have been an acknowledgement that if the court lacked sufficient information to decide whether the applicant in question was entitled to refugee status, it might not have been able to take the decision itself. Here it seems that substitution owing to bias (factor (iii)) would depend on the court being in as good a position (factor (iv)). In *Premier; Mpumalanga*, too, O'Regan J acknowledged the relationship between factors (iii) and (iv):

[A] court should be slow to conclude that there is bias such as to require a court to exercise a discretion particularly where the discretion is one conferred upon a senior member of the executive branch of government.³²

In *Silverstar*, when the decision of the Gauteng Gambling Board (GGB) not to award a gambling licence to Silverstar Development was taken on review, the SCA emphasised that since administrators are 'generally best equipped by the variety of [their] composition, by experience, and [their] access to sources of relevant information and expertise to make the right decision', a court is 'required to recognise its own limitations'.³³ Importantly, this was one of the first cases to suggest – without finding definitively – that factor (iii) (in as good a position) was a kind of threshold consideration for substitution. Heher JA reasoned further that, since 'the court *a quo* was not merely in as good a position as the [GGB] to reach a decision but was faced with the inevitability of a particular outcome if the [GGB] were once again to be called upon fairly to decide the matter', substitution was warranted.³⁴ Although the court did not conclude that factors (i) (foregone conclusion) and (iv) (in as good a position) were necessary preconditions, it did appear to regard the presence of both factors as strengthening the case for substitution.

The court's approach in *Silverstar* was taken further in *Intertrade*, where Plasket J concluded that '[t]he availability of proper and adequate information and the institutional competence of the Court to take the decision' are 'necessary

²⁷ Baxter (note 11 above) at 684 (emphasis added).

²⁸ *Competition Commission* (note 26 above) at para 16.

²⁹ *JCC* (note 9 above) at 76F–77A.

³⁰ *Erf One Six Seven Orchards CC v Greater Johannesburg Metropolitan Council (Johannesburg Administration) and Another* [1998] ZASCA 91, 1999 (1) SA 104 (SCA) ('*Erf 167 Orchards*') at 109F.

³¹ *Harerimana v Chairperson, Refugee Appeal Board and Others* [2013] ZAWCHC 209, 2014 (5) SA 550 (WCC) ('*Harerimana*') at paras 34–42.

³² *Premier; Mpumalanga* (note 9 above) at para 51.

³³ *Silverstar* (note 12 above) at para 29.

³⁴ *Ibid* at para 39.

prerequisites ..., apart from “exceptional circumstances”, for substitution.³⁵ Therefore, despite the court’s finding that there had been an ‘alarming degree of ineptitude, a lack of appreciation of what is required, a lack of judgment, rationality and common sense, and a disturbing contempt for the Constitution and for the people of the province that the Constitution seeks to protect from abuses of public power’,³⁶ Plasket J desisted from ordering substitution. In his view, even gross incompetence and the ‘shameful treatment’ of the aggrieved party over ‘a protracted period’ would not justify this remedy unless the court possessed the necessary information or requisite ‘institutional competence’ to take the decision itself.³⁷ In this regard Plasket J explained that these factors are minimum requirements for rational decision-making.³⁸ The effect of this dictum is that even gross incompetence cannot justify substitution in the absence of factor (iv).

This was also the approach adopted in the subsequent unreported judgment in *M v Minister of Home Affairs*, where the High Court reasoned that provided the court is ‘in a position practically to make the decision’,³⁹ and that one of the factors (i), (ii) or (iv) is also present, substitution will be appropriate. In essence, the court emphasised factor (iii) (in as good a position) as a precondition (in the sense that the court must be able to make the decision itself) and saw the other factors as strengthening the case for substitution.⁴⁰ Bias, for example, would not on its own justify substitution, but would nevertheless support the case for substitution. Notably, however, in contrast to what Plasket J concluded in *Intertrade*, the courts in *Erf 167 Orchards* and *Airoadexpress* had suggested that ‘gross incompetence’ could be an independent basis for substitution.⁴¹

In the earlier judgment of *Livestock*, the court reviewed and set aside a decision of the Livestock and Meat Industries Control Board.⁴² It did so on the basis that the board did not have the power to refuse transfer of registration of a certificate to operate a butchery purely on grounds that the area was ‘adequately and conveniently catered for by the existing butcheries in the area’.⁴³ In deciding whether to substitute, Holmes AJA relied on the following factors: the fact that the board had ‘closed its mind’ to the matter (suggesting that it would not be able to be independent if the matter were remitted); that no other ground (other than the area being adequately catered for) was invoked to justify its refusal (suggesting that the matter was a foregone conclusion); the varying attitude of the board (suggesting possible incompetence); and the delay and frustration that had resulted in a loss of trade profits for the aggrieved party.⁴⁴ The court found that the

³⁵ *Intertrade* (note 9 above) at para 43.

³⁶ *Ibid* at para 5.

³⁷ *Ibid* at para 43.

³⁸ *Ibid*.

³⁹ *M v Minister of Home Affairs* (note 15 above) at para 166.

⁴⁰ *Ibid* at paras 165–166.

⁴¹ *Erf 167 Orchards* (note 30 above) at 109F and *Airoadexpress* (note 20 above) at 680F.

⁴² *Livestock and Meat Industries Control Board v Garda* 1961 (1) SA 342 (A) (*‘Livestock’*).

⁴³ *Ibid* at 348G.

⁴⁴ *Ibid* at 349H–350B.

‘cumulative effect’ of these factors meant that substitution was justified. In effect, none of the factors was decisive, but taken together they merited substitution.⁴⁵

What the above discussion shows is that the courts have taken a varying approach to substitution, essentially oscillating between regarding the factors as individually sufficient, particularly in relation to bias and incompetence;⁴⁶ sufficient in combination with other factors; as necessary prerequisites; or simply factors to be weighed in determining whether to grant substitution. The contradictory way in which the courts have regarded the interplay of factors was highlighted by Khampepe J in *Trencon* as a reason for revisiting the exceptional circumstances test.⁴⁷

To compound this uncertainty, the courts have invoked considerations of fairness in differing ways. As a result, it is unclear whether fairness is the overriding test that the courts apply, a separate factor to be considered alongside the other existing factors, or no more than an underlying value that guides a court’s assessment of these factors in deciding whether exceptional circumstances are present.

B The Role of Fairness

The relevance of fairness has its roots in the much-quoted dictum of Holmes AJA in *Livestock*:

The court has a discretion, to be exercised judicially upon a consideration of the facts of each case, and ..., although the matter will be sent back if there is no reason for not doing so, *in essence it is a question of fairness to both sides*.⁴⁸

Baxter too, relying on *Livestock*, says that fairness is the ‘overriding principle’.⁴⁹ More recently, in *Competition Commission*, the SCA held that ‘[t]here will accordingly be no remittal to the administrative authority in cases where [this] will operate procedurally unfairly to both parties’.⁵⁰ And in *Harerimana*, Davis J said that in deciding whether to order substitution, fairness ‘must be uppermost in the mind of a court’.⁵¹ But is fairness the test that courts apply or is it merely a principle that underlies the concrete criteria discussed in part II.A above? In other words, has the test for substitution been fairness or was it always ‘exceptional circumstances’, with fairness playing a supporting role in the application of the factors?

In *Livestock*, notwithstanding Holmes AJA’s earlier dictum, fairness was not the overriding basis for non-remittal. Rather, it guided the court’s application of various other factors: the fact that the outcome was a foregone conclusion (since there was no lawful basis for the board’s refusal to register the butchery in question); the unfairness that would result from further delay; the unfairness

⁴⁵ Ibid at 350A.

⁴⁶ See further *Vukani Gaming Free State (Pty) Ltd v Chairperson of the Free State Gambling and Racing Board and Others* [2010] ZAFSHC 33 (*Vukani*) at para 26, where the court said that gross incompetence, long delay and inevitability of outcome would individually constitute exceptional circumstances.

⁴⁷ *Trencon* (note 1 above) at para 41.

⁴⁸ *Livestock* (note 42 above) at 349F–H (emphasis added).

⁴⁹ Baxter (note 11 above) at 681, with reference to *Livestock* (note 42 above).

⁵⁰ *Competition Commission* (note 26 above) at para 14.

⁵¹ *Harerimana* (note 31 above) at para 29.

of remitting given the administrator's possible incompetence; and the unfairness of remitting where the board's independence had been brought into question.⁵² Further, in *Competition Commission* the SCA remarked that '[a]ll that can be said is that considerations of fairness may in a given case require the Court to make the decision itself provided it is able to do so'.⁵³ In effect, once the court resolved that it was able to take the decision itself, it considered whether there were other equitable considerations that favoured or militated against substitution. Similarly, in *M v Minister of Home Affairs* the court said:

Fairness to both sides is an important consideration. Sometimes fairness to the applicant can tilt the scale in his or her favour, and considerations of fairness may in a given case require the court to make the decision itself *provided it is able to do so*.⁵⁴

Again, only once factor (iv) (in as good a position) was established did the court consider whether there were other equitable considerations that would support substitution. For example, the court said that the possibility of 'a further lengthy delay' in deciding on the applicant's refugee status and the adverse effect this would have on the applicant's minor daughter tilted the scale in favour of substitution.⁵⁵

In *Intertrade* Plasket J reasoned that fairness underpinned factor (iv) in the sense that without this prerequisite, it would not be fair or rational for the court to take the decision itself.⁵⁶ In the subsequent decision of *Silverstar*, although fairness was seemingly given a much more visible role (the fact that nothing could be 'gained by remittal'⁵⁷ was also relevant to the question of fairness), it was still arguably not elevated to a decisive test. In this regard the court said that since a 'lack of fairness to the Board or the reasonable possibility of prejudice to the public were not probable consequences of non-remittal' (because the outcome was a foregone conclusion) and since there were 'equitable considerations' favouring *Silverstar* (significant delay and the board's stubborn opposition to *Silverstar*), substitution was justified.⁵⁸ This case suggested that fairness underlies factors (i) (foregone conclusion), (ii) (delay) and (iv) (lack of independence) in the sense that where the outcome is certain (together with further equitable considerations), refusing substitution may not be fair. Therefore, even though fairness was given significant attention in this case, its purpose was to inform the question whether there were exceptional circumstances rather than elevate the principle to a decisive test.

In *JCC* the court reasoned that substitution would be justified '[w]here the tribunal or functionary has exhibited bias or incompetence to such a degree that it would be unfair to require the applicant to submit to the same jurisdiction again'.⁵⁹ Here the principle underpinned factor (ii) and was concerned with the

⁵² *Livestock* (note 42 above) at 349H–350B.

⁵³ *Competition Commission* (note 26 above) at para 15 (emphasis added).

⁵⁴ *M v Minister of Home Affairs* (note 15 above) at para 165 (emphasis added).

⁵⁵ *Ibid* at paras 175–176; *Vukani* (note 46 above) at para 51.

⁵⁶ *Intertrade* (note 5 above) at para 43.

⁵⁷ *Silverstar* (note 12 above) at para 40.

⁵⁸ *Ibid*.

⁵⁹ *JCC* (note 9 above) at 76F–H.

unfairness of remitting to a biased administrator. This also seems to have been the approach in *Harerimana*, where Davis J reasoned that

[t]o remit the appeal to the body, which in so determined a fashion has resisted this review application, would, at the very least, be unfair. The conduct of [the board] promotes a perception that it is a body which is now biased and not sufficiently independent in respect of [the] applicant's case. It would thus be unfair for it to be required to reconsider its own decision, that it obviously considers to be correct, both in substance and procedurally.⁶⁰

While Hoexter is correct that '[f]airness to both sides has always been and will most certainly remain an important consideration',⁶¹ the courts have not clearly articulated what role fairness actually plays in deciding whether to substitute. That said, rather than an overarching test for substitution, considerations of fairness appear to have played a supporting role in relation to factors (i)–(iv), some favouring substitution (such as the fact that the court is in as good a position to take the decision itself) and others cutting against it (such as the fact that the outcome is not foregone or the fact that further delay occasioned by remittal would not be prejudicial to the aggrieved person).

But in *Trencon*, rather than refining our understanding of fairness in the context of a substitution order, the decision has, as we shall see, left it unclear whether fairness is now the decisive test for determining whether substitution is warranted; a separate factor that must be weighed against factors (i)–(iv);⁶² or whether it remains an underlying value that reveals which factors are relevant and how they should be weighed. And more seriously, in its drive for flexibility and preference for ill-defined conceptions of fairness, the Court has failed to provide sufficient direction to lower courts and litigants on how to exercise the discretion in question. The flexibility required by the Constitution (in this context, determining the 'just and equitable' remedy) does not sanction murky judicial reasoning, nor does it permit judges to shy away from laying down clear principles of law.⁶³

The upshot is that a guided approach to substitution is required, one which, for example, clarifies the existence or otherwise of any strict requirements for substitution; the relative weight to be attached to each of the existing factors; and ultimately, the role of fairness in the overall decision to substitute or remit. Guidance is not the same as strait-jacketing the discretion in question. It means

⁶⁰ *Harerima* (note 31 above) at para 32.

⁶¹ C Hoexter *Administrative Law in South Africa* (2nd Edition, 2012) 553.

⁶² See, eg, *Reizis* (note 22 above) at para 39, where the court applied fairness as an independent factor alongside factors of foregone conclusion, bias and incompetence, and the prospect of further delay.

⁶³ See S Woolman 'The Amazing, Vanishing Bill of Rights' (2007) 124 *South African Law Journal* 762, 788–792 and especially at 789, where he quotes Justice O'Regan: 'As a member of the bench, I am required to issue a judgment; and in that judgment, I am required to lay down a rule of law that binds both the parties before the court and South African society as a whole; however, if laying down a rule of law makes me, ipso facto, a positivist, then a positivist of some stripe my office commits me to be.' (Justice K O'Regan 'Dignity, Application and the Rule of Law', a paper presented at the *Conference on Dignity and the Jurisprudence of Laurie Ackermann* University of Cape Town, 27 July 2007.) On the value of legal rules, see in general A Fagan & F Schauer 'Rules – in Law and Elsewhere' in F Schauer (ed) *Thinking Like a Lawyer: A New Introduction to Legal Reasoning* (2009) Chapter 2, especially at 35.

that ‘courts ... can – and ought to – generate rules of law that are determinate enough to guide the behaviour of state and non-state actor[s]’.⁶⁴

III *TRENCON: THE FACTS*

On 18 May 2012 the Industrial Development Corporation (IDC) invited tenders from building contractors for the upgrade of its head office in Sandton, Johannesburg. The tender followed a two-step trajectory, beginning with a pre-qualification phase (to identify the best proposals based on technical capability, environmental management, personnel capabilities, financial standing and litigation history) followed by a competitive-bidding phase (involving actual bidding and an evaluation based on considerations such as price and broad-based black economic empowerment). The process engaged the expertise of various procurement bodies at the IDC, as well as independent consultants.

In the first phase the Request for Proposals stipulated that the closing date for submission would be Monday 4 June 2012 at noon. Those contractors who were shortlisted in the first round would be eligible to participate in the competitive-bidding phase that followed. Trencon Construction (Pty) Ltd, Basil Read (Pty) Ltd and five other contractors were shortlisted and invited to tender on the basis that the site handover would be on 6 September 2012. The second phase involved a multistage process of evaluation, assessment and approval by various procurement bodies at the IDC, culminating in a final decision on the award of the tender by its Executive Committee (Exco).

Trencon tendered a price of R 117 112 972,21 provided, crucially, that the site handover date of 6 September 2012 remained unchanged. During the evaluation process, the consultants sought clarification from the bidding parties on the cost implications if the site handover were delayed until 1 October 2012, or even beyond that date. In response, Trencon advised that it would charge a 6 per cent escalation fee (approximately R 315 000,00 excluding VAT), whereas Basil Read indicated that its price would remain fixed notwithstanding the delay. When the various bids were evaluated, all the IDC’s in-house procurement bodies, together with the consultants, recommended that Trencon be appointed as the preferred contractor (subject to certain conditions, such as a revised contract value). Importantly, even with the price increase, Trencon’s offering scored the highest points on both price and empowerment. In spite of this, when the final decision went before the Exco, it resolved to appoint Basil Read. Trencon’s tender had, in its view, failed to maintain a fixed price for the 120-day period as required by the tender conditions and was therefore non-responsive.

Trencon approached the High Court to review the IDC’s decision primarily on the basis that the latter had misunderstood the tender provisions relating to contract price adjustment by failing to differentiate between CPAP Adjustments (increases in the costs of labour, materials, plant and goods) and Default Adjustments (increases arising from other delays such as site handover). While the former costs were required to remain fixed for 120 days, price increases arising from the latter were not prohibited. The High Court accepted this distinction,

⁶⁴ Woolman (note 63 above) at 789.

holding that the IDC had committed a material error of law in concluding that the tender documents prohibited any price escalations in relation to delays in site handover.⁶⁵ Trencon's bid was thus found to be compliant and responsive, and the award of the tender to Basil Read was set aside.

However, Mothle J went further, finding that there were exceptional circumstances warranting substitution: Trencon's bid had scored the highest points throughout the process; it had been favoured as the preferred bidder by all of the IDC's procurement bodies; and the IDC had been unable to present any evidence why the tender should not be awarded to Trencon or why the court would be justified in cancelling the contract, ordering that the process start afresh, and thus further delaying implementation of the project.⁶⁶ Regarding the alternative argument that the matter be remitted to the IDC with instructions to award the tender to Trencon, the court reasoned essentially that it 'would make no difference' if it were to take the decision itself and that the decision, barring the material error of law, was a foregone conclusion.⁶⁷ Mothle J also remarked that the court was 'qualified' to take the decision itself, and that, considering the 'substantial' amount of public funds involved and the fact that further delay would 'cause unjustifiable prejudice to Trencon, the IDC and National Treasury', there was no reason, given the 'urgency of the matter', to refer the decision back to the IDC.⁶⁸ In his reasoning, Mothle J prioritised factor (i) (foregone conclusion) and seemed to regard factors such as delay (factor (ii)), that the court was qualified to take the decision (factor (iv)) and that the tender implicated public money, as strengthening the case for substitution.

On appeal, the SCA accepted that the IDC had made a material error of law, but did not agree that substitution was warranted.⁶⁹ This was notwithstanding its acceptance that the only reason for Trencon's non-appointment, which the IDC also conceded, was the said error.⁷⁰ The SCA said that the High Court had 'overlooked the fact that the IDC was not obliged to award the tender to the lowest bidder or at all', for which reason the outcome was not a foregone conclusion.⁷¹ It further criticised the High Court for failing to give proper consideration to the separation of powers as well as the question of delay.⁷² In the latter regard Maya JA reasoned that over two years had elapsed since the beginning of the tender process, which meant that the information before the court, and the basis upon which the tenders had been evaluated, was outdated. The lower court's order thus failed to consider 'unavoidable supervening circumstances such as price increases'.⁷³ In essence, the SCA found that none of factors (i)–(iv) had been present and that substitution could not, therefore, be justified.

⁶⁵ *Trencon Construction (Pty) Ltd v Industrial Development Corporation of South Africa Ltd and Another* [2013] ZAGPPHC 147, 2013 JDR 1267 (GNP) ('Trencon HC') at paras 28–31.

⁶⁶ *Ibid* at para 53.

⁶⁷ *Ibid*.

⁶⁸ *Ibid*.

⁶⁹ *Industrial Development Corporation of South Africa Ltd v Trencon Construction (Pty) Ltd and Another* [2014] ZASCA 163, 2014 (4) All SA 561 (SCA), 2014 JDR 2154 (SCA) ('Trencon SCA').

⁷⁰ *Ibid* at para 11.

⁷¹ *Ibid* at para 18.

⁷² *Ibid*.

⁷³ *Ibid* at para 19.

IV EXCEPTIONAL CIRCUMSTANCES IN *TRENCON*: THE MISSED OPPORTUNITY

Owing to the lack of guidelines in s 8(1)(c)(ii)(aa) of the PAJA, Khampepe J said that it was of ‘great import’ that the test for exceptional circumstances be revisited.⁷⁴ But rather than sharpening the existing approach, the Court has obfuscated it through a rather complicated weighing exercise that seems to turn, ultimately, on what is fair:

[G]iven the doctrine of separation of powers, in conducting [the] enquiry there are certain factors that should inevitably hold greater weight. The first is whether a court is in as good a position as the administrator to make the decision. The second is whether the decision of an administrator is a foregone conclusion. These two factors must be considered cumulatively. Thereafter, a court should still consider other relevant factors. These may include delay, bias or the incompetence of an administrator. The ultimate consideration is whether a substitution order is just and equitable. This will involve a consideration of fairness to all implicated parties. It is prudent to emphasise that the exceptional circumstances enquiry requires an examination of each matter on a case-by-case basis that accounts for all relevant facts and circumstances.⁷⁵

In parts 1–3 below I propose to unpack and ultimately critique the Court’s stated test based on the following broad structure and using somewhat different numbering:

1 *Separation of Powers*

- (a) Is the court in as good a position as the administrator to make the decision?
And if so,
- (b) Is the decision a foregone conclusion?

2 *Other Relevant Factors*

- (c) If (a) and (b) are answered affirmatively, are there further considerations such as delay, good faith, bias or incompetence that favour substitution?

3 *Fairness*

- (d) Notwithstanding the existence of (a), (b) and (c), would substitution be just and equitable in the circumstances?

1 Separation of Powers

In the earlier stages of the litigation, the SCA criticised the High Court for failing to balance the substitution remedy against the requirements of the separation of powers and of judicial deference.⁷⁶ This moved Khampepe J to prioritise considerations (a) and (b). In particular, she reasoned that the test should be informed ‘not only by the deference courts have to afford an administrator but also by the appreciation that courts are ordinarily not vested with the skills and expertise required of an administrator’.⁷⁷

⁷⁴ *Trencon* (note 1 above) at para 41.

⁷⁵ *Ibid* at para 47.

⁷⁶ *Trencon* SCA (note 69 above) at para 18.

⁷⁷ *Trencon* (note 1 above) at para 43.

While there is nothing objectionable in principle about an approach that emphasises these factors, which is arguably what earlier cases endeavoured to do,⁷⁸ the real difficulty arises from the disconnect between the chronological scheme that the Court proposes and the particular wording of the judgment. Specifically, Khampepe J explains that only '[o]nce a court has established that it is in as good a position as the administrator' will it be '*competent to enquire* into whether the decision of the administrator is a foregone conclusion',⁷⁹ and that 'thereafter'⁸⁰ it should consider other relevant factors. This appears to impose a chronological trajectory that a court should follow in determining whether the case is an exceptional one. But then, confusingly, Khampepe J says that factors (a) and (b) 'inevitably hold greater weight' and that they must be considered 'cumulatively'.⁸¹

Though in her evaluation of the case Lauren Kohn recognises that it is not 'entirely clear' from the judgment whether factors (a) and (b) are 'strict requirements' or simply requirements that weigh more heavily, she argues that on a purposive interpretation of the judgment, the Court intended these factors to be strictly present, particularly given Khampepe J's emphasis on accommodating the separation of powers within the rubric of the test.⁸² But my argument is that it is simply unclear from the wording of the judgment – and even the scheme proposed by the judgment as a whole – whether factors (a) and (b) are intended to be preconditions for substitution (in the sense that a box must be ticked in each case before the factors listed in (c) can be considered), or whether these factors must merely be weighted more heavily against other factors, such as those in (c) and ultimately fairness in (d). That said, there is a strong argument that at least factor (a) should be treated as a strict requirement for substitution. This, as Plasket J explained in *Intertrade*, is because it would almost never be rational or fair to order substitution where a court is not in as good a position as the administrator to take the decision itself.⁸³ In the end, if the Court intended to impose preconditions for substitution – and in doing so, to limit a court's discretion in ordering the remedy – then it ought to have said so openly and clearly.

Turning to the distinction between (a) and (b), Khampepe J recognises the difficulty of drawing a bright-line divide between these two factors, explaining that they are 'interrelated and interdependent'.⁸⁴ Nevertheless, she says that the difference lies in the fact that factor (a) speaks to the nature of the decision: '[E]ven where the administrator has applied its skills and expertise and a court has all the relevant information before it, the *nature of the decision* may dictate that a court defer to the administrator', which she reasons is 'typical in instances of policy-laden and polycentric decisions'.⁸⁵ Conversely, in circumstances where the decision 'is not polycentric and is guided by particular rules or by legislation',

⁷⁸ See, eg, *Intertrade* (note 5 above) at para 43.

⁷⁹ *Trencon* (note 1 above) at para 49 (emphasis added).

⁸⁰ *Ibid* at para 47.

⁸¹ *Ibid*.

⁸² Kohn (note 6 above) at 103.

⁸³ *Intertrade* (note 5 above) at para 43.

⁸⁴ *Trencon* (note 1 above) at para 50.

⁸⁵ *Ibid* (emphasis added).

the decision may warrant a lesser degree of deference to the administrator.⁸⁶ In effect, Khampepe J seems to be saying that to satisfy requirement (a) a court must be both qualified to make the decision (in the sense that it has the necessary information before it) and institutionally competent to do so, in the sense that the decision must not be of a policy-laden or polycentric kind.

While the High Court failed to provide any explanation as to why it thought it was qualified to take the decision itself,⁸⁷ Khampepe J introduced two helpful substantive requirements to justify her conclusion that the court was in fact qualified: (1) The administrator's expertise must not be required at all (where, for example, the decision is of a judicial nature) or, (2) if it is required, it must not be required any longer (where, for example, the administrative process is at a stage where any specialised knowledge has already been applied). In the case of both (1) and (2) the court must also have all the pertinent information before it in order to take the decision itself. Khampepe J concluded that (2) was satisfied for the following reasons: The material error of law had occurred in the final stages of the procurement process; 'all technical components of the process' had already been undertaken by the various IDC bodies; and all that was left was for Exco to approve the recommendations.⁸⁸ The Court had also had 'the benefit of the record, with all the pertinent information and recommendations that would have been before Exco'.⁸⁹ Moreover, the Court said that the IDC had failed to explain 'how its administrative expertise could come into play at this point or on what basis it could decide differently'.⁹⁰

For the above reasons, Khampepe J concluded that the court was as qualified as the IDC to award the tender to Trencon. She did not, however, consider whether the court was institutionally competent to make that order. If factor (a) is in fact a strict requirement, as the Court can be read to have suggested, then apart from being qualified to order substitution, Khampepe J ought, in addition, to have considered whether taking the decision itself would have given rise to issues of polycentricity. In explaining why polycentric matters are generally ill-suited for consideration by courts, Fuller referred to the multitude of intersecting interests and issues to which such decisions can give rise.⁹¹ In this case, for instance, a relevant consideration may have been the fact that a partially fulfilled contract would have to be terminated and replaced with a new agreement with considerable financial implications for both the IDC and the public purse.

Turning to factor (b), foregone conclusion, the Court said that this is satisfied where there is 'only one proper outcome of the exercise of an administrator's discretion and "it would merely be a waste of time to order the [administrator] to reconsider the matter"'⁹² or, put differently, 'remittal would serve no purpose' since there would be no 'discretion left to exercise'.⁹³ In concluding that the

⁸⁶ Ibid at para 49.

⁸⁷ *Trencon* HC (note 65 above) at para 53.

⁸⁸ *Trencon* (note 1 above) at para 57.

⁸⁹ Ibid at para 58.

⁹⁰ Ibid.

⁹¹ See L Fuller 'The Forms and Limits of Adjudication' (1978–79) 92 *Harvard Law Review* 353.

⁹² *Trencon* (note 1 above) at para 49, quoting from *JCC* (note 9 above) at 76E–F.

⁹³ *Trencon* (note 1 above) at para 59.

outcome of the tender process was foregone, the court relied on the fact that Trencon had scored the highest points throughout the process (even following the proposed price escalation); that all of the IDC's in-house bodies as well as its external experts had recommended that it be awarded the tender; and that it was common cause that Trencon's non-appointment could only be attributed to Exco's material error of law. In response to the IDC's contention that it retained a discretion not to award the tender to the highest points-earner, or not to award it at all, the Court reasoned that, irrespective of the validity of any other bid, the question was whether the IDC was bound to award the tender to Trencon. And, in the absence of 'objective criteria or justifiable reasons',⁹⁴ it concluded that the IDC was in fact bound, especially because Trencon was the highest points-earner and Exco had not cited any other reason (besides its material error of law) for its decision.⁹⁵

What is of crucial significance in relation to this factor is Khampepe J's emphatic view that a court will be competent to enquire whether this factor has been met only after it has made a positive finding in (a). In her assessment, 'there can never be a foregone conclusion unless a court is in as good a position as the administrator'.⁹⁶ She also explained that it would be more difficult to establish the inevitability of a particular outcome where the administrator's unique expertise and experience have not been adequately applied to the matter.⁹⁷ This statement demonstrates a significant development from the traditional approach and does provide some clarity, since establishing that a particular outcome is a foregone conclusion is now dependent on a court's having the requisite institutional expertise and competence to decide the issue.

2 Other Relevant Factors

While the cases reflecting the traditional approach did not distinguish between different forms of delay, Khampepe J draws a distinction between a delay resulting in a change of circumstances (or pre-judgment delay) and delay occasioned by remittal. This seems to be what she had in mind when she said that delay 'could cut both ways'.⁹⁸ Regarding the second form, if the further delay arising from remittal would be prejudicial to the affected party, who has already waited a very long time for a decision, then substitution may be appropriate. This is in contrast to the first form, where delay may, for instance, favour remittal if the administrator's needs have changed fundamentally; the circumstances have been drastically altered; or if further delay would have a negative impact on the public purse (or, presumably, on those depending on the services being procured). This consideration is particularly relevant to the procurement context, where tenders depend very much on the needs of the administration at a particular time. However, the Court's suggestion that a drastic or fundamental change in circumstances occasioned by delay would favour remittal may be read to suggest

⁹⁴ Ibid at para 63.

⁹⁵ Ibid at para 65.

⁹⁶ Ibid at para 50.

⁹⁷ Ibid at para 49.

⁹⁸ Ibid at para 51.

that a lesser change would not. In response, I argue that even a minor change may favour remittal depending on how this factor is weighted against the other relevant factors.

In relation to delay occasioned by the litigation process – a form of pre-judgment delay – the Court was emphatic: ‘What must be stressed is that [such a delay] should not easily clout the court’s decision in reaching a just and equitable remedy’⁹⁹ and that, particularly in appeals, ‘delay is inevitable’.¹⁰⁰ Moreover, Khampepe J cautioned that if a litigation delay were always to work against substitution, it would have the undesirable consequence of encouraging state parties to engage in protracted litigation so as to avoid substitution.¹⁰¹ In essence, she said that delay caused by the appeal process will almost never work against substitution owing to the rule that ‘an appeal should be decided on the facts that existed when the original decision was made’.¹⁰² In support of this, the Court relied on a general statement made by Froneman J in *Billiton Aluminium*.¹⁰³ But even in that judgment it was recognised that ‘[t]his is not an inflexible rule and after-the-fact evidence may be admitted ... in “exceptional cases that cry out for the reception of post-judgment facts”’.¹⁰⁴ And, more pointedly, surely this rule cannot extend to the remedy that is just and equitable. If facts on the ground have changed, of course that is relevant to what relief is going to be appropriate in the circumstances.¹⁰⁵ In particular, if the tender is now not needed at all, why must the government procure something that it no longer needs? Or if the needs of government have changed or the market landscape could provide more competitive bids, why should it be bound by the existing tender and the present bidders? Khampepe J said that the public purse matters as a reason in favour of substitution. Why then could it not also work against substitution, as these examples suggest that it should?

Khampepe J went on to say that even if reliance could be placed on the delay occasioned by changed circumstances in this case (which, she added, was not possible since the IDC had not adduced any evidence to this effect),¹⁰⁶ a mere change in price was insufficient to establish a change in circumstances.¹⁰⁷ It is not entirely clear what she means by a change in price, but presumably she is referring to the price bidders would be willing to charge. Wishing to take the distinction between public and private law seriously, Khampepe J then reasons that contract-price adjustment is the subject of ‘ordinary contractual negotiations’ between the parties and cannot, therefore, have a bearing on the appropriate remedy in public

⁹⁹ Ibid at para 52.

¹⁰⁰ Ibid at para 53.

¹⁰¹ Ibid.

¹⁰² Ibid at para 52.

¹⁰³ *Billiton Aluminium SA Ltd t/a Hillside Aluminium v Khanyile and Others* [2010] ZACC 3, (2010) 31 ILJ 273 (CC), 2010 (5) BLLR 465 (CC) (‘*Billiton Aluminium*’) at para 35.

¹⁰⁴ Ibid at para 36, quoting the words of Comrie J in *Van Eeden v Van Eeden* 1999 (2) SA 448 (C) at 453A.

¹⁰⁵ Section 38 of the Constitution states that ‘[a]nyone listed in this section has the right to approach a competent court, alleging that a right in the Bill of Rights has been infringed or threatened, and the court may grant appropriate relief’.

¹⁰⁶ *Trencon* (note 1 above) at para 73.

¹⁰⁷ Ibid at para 72.

law.¹⁰⁸ But logically, the fact that a bidder may no longer be willing to charge the price at which it tendered initially must necessarily impact on a court's decision to order substitution, irrespective of the area of law it is apparently subject to.

Traditionally, factors such as bias or incompetence have sometimes been treated as sufficient on their own to warrant substitution and sometimes not.¹⁰⁹ *Trencon* perpetuates this uncertainty. In relation to bias or incompetence, consistent with the traditional view, the Court says that where it would be unfair to require the applicant to resubmit itself to the administrator's jurisdiction, this 'would weigh heavily in favour of substitution'.¹¹⁰ Good faith is a further factor considered in *Trencon*, where Khampepe J explains that 'the fact that the IDC acted in good faith when it was moved by a material error of law should be a *strong consideration*'.¹¹¹ However, stating that these factors weigh heavily or are strong considerations is not guidance enough, since it remains unclear from the scheme of the judgment whether any of the factors under (c) can be considered in the absence of establishing (a) and (b). On the one hand, if (a) and (b) are interpreted as strict requirements then the factors in (c) will never be separate grounds upon which substitution can be granted. On the other hand, if (a) and (b) are not strict requirements, but only have to be weighted more heavily, it seems possible that the presence of any one of these factors could justify substitution – assuming that this is what fairness demands.

3 Fairness

Relying on the dictum in *Livestock*, Khampepe J said that even before the PAJA, 'courts almost invariably considered the notion of fairness'.¹¹² In articulating a role for fairness, she located the power to substitute in s 8(1)(c)(ii)(aa) of the PAJA in the context of a court's wider discretion to make 'any order that is just and equitable' in terms of s 8(1).¹¹³ The effect, as Khampepe J explained, is that 'even where there are exceptional circumstances, a court must be satisfied that it would be just and equitable to grant an order of substitution'.¹¹⁴ Therefore, in addition to exceptional circumstances, a court would still be required to satisfy itself that substitution would be fair to all parties. The suggestion here is that fairness is essentially a separate factor to be taken into account over and above the existing factors. But later on Khampepe J wrote that '[t]he *ultimate consideration* is whether a substitution order is just and equitable'¹¹⁵ and went on to explain (by example) that 'having regard to the notion of fairness, a court may still substitute even where there is no instance of bias or incompetence'.¹¹⁶ Taken to their logical conclusion, these latter statements could be read to suggest that what matters

¹⁰⁸ Ibid at para 75.

¹⁰⁹ Compare, eg, *JCC* (note 9 above) at 76F–77B and *Competition Commission* (note 26 above) at paras 17–18 with *Intertrade* (note 5 above) at para 43.

¹¹⁰ *Trencon* (note 1 above) at para 54.

¹¹¹ Ibid at para 77 (emphasis added).

¹¹² Ibid at para 36.

¹¹³ Ibid at para 35.

¹¹⁴ Ibid.

¹¹⁵ Ibid at para 47 (emphasis added).

¹¹⁶ Ibid at para 54.

ultimately is fairness, or simply that fairness (not exceptional circumstances) is the test that courts should apply.

Despite this somewhat confusing formulation, a constructive reading of the judgment suggests that it was probably not Khampepe J's intention to elevate fairness to being the overriding test or to treat it as a separate factor, but rather to introduce a degree of flexibility to ensure that courts do not formalistically stick to the established factors by, for example, engaging in a rigid box-ticking exercise in deciding whether substitution is justified. My view is that it was not necessary for the Court to use fairness in this way: the existing factors have never been exhaustive and it has always been open to the courts to add to the list in light of the facts and circumstances of a particular case. The flexibility that Khampepe J is calling for is therefore inherent in the discretionary power itself and need not be derived from the principle of fairness. The crucial point is that the established factors are, in fact, concrete manifestations of equity, some motivate for substitution (such as the fact that the outcome is forgone) while others cut against it (such as where the delay has meant that the existing contract is near completion). In effect, the function of fairness is to inform the question whether exceptional circumstances are present. Put differently, it is the underpinning or normative basis of the existing factors. It is certainly not, and has never been, a separate, stand-alone reason for substitution.

In the subsequent case of *Westinghouse*,¹¹⁷ the SCA revisited the *Trencon* test. In this case, following the finding that Eskom had unlawfully awarded a tender to Areva for the replacement of steam generators at the Koeberg nuclear power station in the Western Cape, the SCA was asked to substitute Areva with another contractor, Westinghouse. Notwithstanding the finding that Eskom had acted unlawfully in relying on considerations not contained in the bid criteria, Lewis JA held that it would not be equitable to award substitution for the following reasons: Eskom should still be given an opportunity to rerun the tender process and to decide whether the criteria it had considered was in fact vital;¹¹⁸ further delay resulting from substitution would be undesirable given that the work under the existing contract had already commenced and that there was a deadline looming;¹¹⁹ and the award of the tender to Westinghouse was not clearly a foregone conclusion, especially if these additional considerations (that Eskom deemed vital) were taken into account.¹²⁰ In deciding whether substitution was justified Lewis JA, applying the *Trencon* test, did not treat fairness as a separate factor or a decisive test. Rather, these concrete factors (with their basis in fairness) informed, and ultimately justified, the court's decision to remit rather than substitute.

The Court's approach to fairness in *Trencon* also raises important questions about its repeated preference for flexibility at the expense of developing some

¹¹⁷ *Westinghouse Electric Belgium SA v Eskom Holdings (SOC) Ltd and Another* [2015] ZASCA 208, 2016 (3) SA 1 (SCA), 2016 (1) All SA 483 (SCA) ('*Westinghouse*'). When this judgment went on appeal to the Constitutional Court in *Areva NP Incorporated in France v Eskom Holdings Soc Ltd and Others* [2016] ZACC 51, the issue of substitution was not revisited owing to the Court's dismissal of the matter on a narrow procedural basis.

¹¹⁸ *Westinghouse* (note 117 above) at para 76.

¹¹⁹ *Ibid* at para 77.

¹²⁰ *Ibid* at para 78.

real law.¹²¹ While it is easy (and at times justified) to criticise rules as being formalistic and characteristic of a different age of adjudication unsuited to the kind of transformative adjudication that our Constitution demands, the fact remains that there is significant value in rule-based reasoning: it is helpful for litigants and courts trying to understand and apply the law. This can provide a safety net against arbitrary and unjust decision-making by guiding the exercise of judicial discretion, and has the real benefit of truncating litigation by encouraging predictability in legal outcomes.¹²² A good example of this kind of reasoning is the *Intertrade* judgment, which essentially found that no matter how egregiously an administrator has behaved, it can never be fair or rational for a court to order substitution where it lacks the necessary information or requisite institutional competence to take the decision itself.¹²³ In effect, the judgment stands for the proposition that in the absence of factor (a), ordering substitution would, in general, not be justified. The purpose of laying down this rule was not to impose rigidity on a court's discretionary powers but to provide direction on how this particular discretion ought to be exercised.

V CONCLUDING REMARKS

This note has levelled two main criticisms against the exceptional-circumstances enquiry in *Trencon*. The first relates to the absence of clarity as to the nature and content of (and also the relationship between) the various factors that underlie the test, both under the traditional approach and as they were interpreted and applied in *Trencon*. The second relates to the uncertainty occasioned by the Court's

¹²¹ See, eg, *Lee v Minister for Correctional Services* [2012] ZACC 30, 2013 (2) SA 144 (CC), 2013 (2) BCLR 129 (CC), where the Court introduced a flexible standard for assessing factual causation in the delictual context. In essence, it reasoned that where a strict application of the common-law test for factual causation would lead to an injustice, it should be applied flexibly. Moreover, in *South African Police Service v Solidarity obo Barnard* [2014] ZACC 23, 2014 (6) SA 123 (CC), 2014 (10) BCLR 1195 (CC) a minority of the Court concluded that the appropriate standard for assessing the lawfulness of the implementation of a valid affirmative action measure was fairness. In a contractual context, in *Barkhuizen v Napier* [2007] ZACC 5, 2007 (5) SA 323 (CC), 2007 (7) BCLR 691 (CC) Ngcobo J found that a time-bar clause in a contract could be declared contrary to public policy on grounds that it was unreasonable or unfair. This dictum has given rise to the impression that fairness is now the standard for not enforcing promises in contracts, as evidenced by the subsequent approach in *Botba and Another v Rich NO and Others* [2014] ZACC 11, 2014 (4) SA 124 (CC), 2014 (7) BCLR 741 (CC), where the Court concluded that '[t]o the extent that the rigid application of the principle of reciprocity may ... lead to injustice, our law of contract, based as it is on the principle of good faith, contains the necessary flexibility to ensure fairness' (para 45). Again, in *Cool Ideas 1186 CC v Hubbard and Another* [2014] ZACC 16, 2014 (4) SA 474 (CC), 2014 (8) BCLR 869 (CC), which concerned the enforceability of a private arbitration award in contravention of a statute, Froneman J resorted in his minority judgment to the standard of fairness as a reason for enforcing the arbitration award. He reasoned that '[p]ublic policy in the interpretation, application and enforcement of contracts embraces issues of fairness. Fairness is one of the core values of our constitutional order. When the enforcement of arbitration awards on the basis of public policy is at stake, fairness lies at the heart of the enquiry, not at its periphery' (at para 126).

¹²² See F Schauer 'Formalism' (1998) 97 (4) *Yale Law Journal* 509. See further JC Froneman 'Legal Reasoning and Legal Culture: Our "Vision" of Law' (2005) 16 *Stellenbosch Law Review* 3 and G Quinot 'Substantive Reasoning in Administrative-Law Adjudication' (2010) III *Constitutional Court Review* 111. Both Froneman and Quinot recognise the value, and necessity, of formal (not formalistic) reasoning in adjudication.

¹²³ *Intertrade* (note 5 above) at para 43.

resort to fairness as a flexibility-enhancing mechanism in the overall substitution enquiry.

In justifying the Court's approach, Khampepe J explained that it 'accords with the *flexibility* embedded in the notion of what is just and equitable'.¹²⁴ But in fixating on flexibility, the Court has compromised on some much-needed clarity in this area and the importance of giving guidance to courts and litigants on how this discretion ought to be exercised. And while there is an inevitable tension between certainty, on the one hand, and the flexibility of more variable standards on the other, a legal system simply cannot function without a minimum level of certainty.¹²⁵ Woolman says the following in this regard:

An approach to constitutional adjudication that makes it difficult for lower court judges, lawyers, government officials and citizens to discern, with *some degree of certainty*, how the basic law is going to be applied, and to know, with *some degree of certainty*, that the basic law is going to be applied equally, constitutes a paradigmatic violation of the rule of law.¹²⁶

For these reasons – the lack of clarity as to the interplay between the factors and Khampepe J's treatment of fairness – I cannot agree with my colleague, Lauren Kohn, that the Court achieved the desired balance between certainty and flexibility. In the end, and whatever the state of public procurement may be in South Africa, the courts need a principled, guided basis for deciding whether to substitute. *Trencon* missed its chance.

¹²⁴ *Trencon* (note 1 above) at para 55 (emphasis added).

¹²⁵ See *Mighty Solutions t/a Orlando Service Station v Engen Petroleum Ltd and Another* [2015] ZACC 34, 2016 (1) SA 621 (CC), 2016 (1) BCLR 28 (CC) at para 37 and *Ferreira v Levin NO and Others; Vryenhoek and Others v Powell NO and Others* [1995] ZACC 13, 1996 (1) SA 984 (CC), 1996 (1) BCLR 1 (CC) at para 26, where the court acknowledges the centrality of legal certainty in our constitutional state.

¹²⁶ Woolman (note 63 above) at 763 (emphasis added).

Affirmative Action

Dignity and Equality in *Barnard*

Samantha Vice*

I INTRODUCTION

At the fall of apartheid in 1994, white South Africans were reassured by the Constitution¹ of the democratic South Africa that their rights as a minority would be respected and that they would be full and valued citizens in the hopeful new country. The preamble of the Constitution promises to ‘establish a society based on democratic values, social justice and fundamental human rights’; s 9 declares that no person can be discriminated against on grounds of race (amongst others); s 25 promises to respect property rights. Despite the Truth and Reconciliation Commission, few white people were explicitly tried and punished for apartheid crimes, and for many, the transition to the ANC-led democracy brought less upheaval than they had expected.

Fears still linger, however, as whites face a reality in which they have to compete with blacks for jobs, and in which their privileged position – which largely remains – never feels quite morally or politically stable. The constitutional case *South African Police Service v Solidarity obo Barnard*² brought one of these fears to the surface, that they would be overlooked for jobs and deprived of opportunities, despite the Constitution’s commitment to protecting their dignity and equality. Here was a case that seemed to make the rumours and anxieties of parts of the white world justified and to reveal their real, undervalued position in the country, whatever its non-racial pretensions. From the perspective of some disaffected whites, ‘affirmative action’ measures are really ‘reverse discrimination’, racial bias directed against them in the face of the reassurances of the Constitution. From the perspective of many whites, therefore, *Barnard* tests the sincerity of the government’s promises to protect them from ‘majoritarian retribution’³ and to consider them equal citizens. It also tests the force of the Constitution. For many South Africans of all colours, the case places under the spotlight how far the courts would permit the government to implement restitutionary measures to improve equality, and whether the Constitution can act as ‘a restraining influence on excessive consolidation of political power’, as Samuel Issacharoff notes.⁴ The justices of the Court were faced, it seemed, with a direct conflict between respecting the dignity and moral equality of a particular person – in this case,

* Professor, Department of Philosophy, University of the Witwatersrand. The author acknowledges, with thanks, the helpful comments of the editors and two anonymous reviewers for this journal.

¹ Constitution of the Republic of South Africa, 1996.

² *South African Police Service v Solidarity obo Barnard* [2014] ZACC 23, 2014 (6) SA 123 (CC), 2014 (10) BCLR 1195 (CC) (*‘Barnard’*).

³ S Issacharoff ‘The Democratic Risk to Democratic Transitions’ (2013) 5 *Constitutional Court Review* 1, 5.

⁴ *Ibid* at 3.

Renate Barnard, a white woman passed over for promotion – and respecting attempts by government agencies to have a more representative work-force, as part of ‘realising’ socio-economic equality. The justices saw this as a conflict, but also tried heroically to come to a decision that both respected the mandate to address inequality and that was not unfairly discriminatory against particular, valuable, individual persons.

Barnard raises many issues: moral, political, jurisprudential. I come to it as a philosopher not a legal scholar, and I will largely explore two (sets of) themes raised throughout the judgments. The first is how the judges handled the (apparent) conflict in this case between the rights of individuals and the rights of people as members of previously disadvantaged groups; or, how they understood and attempted to resolve the conflict between the dignity and moral equality of individual persons, and the ideal of realising social and economic equality for those disadvantaged by apartheid. Exploring this issue requires examining how the judges understood the notions of dignity, equality and fairness. The second is how white people might understand and relate to the judgment against Ms Barnard, and so to their own position in South Africa.⁵

The justification for examining the first set of interests is unproblematic: we rely on the Constitutional Court to interpret the guiding ideals of the Constitution and to come to fair decisions in hard cases. If its judgments reveal the justices’ own confusions or ambiguities in engaging with crucial concepts, then they are on shaky grounds and do not offer clear guidance for future cases. Given the inevitability (as I shall argue) of conflict between the ideals set out in the Constitution, justices in future cases will need guidance and convergence on, at least, the meaning of their principles and values. I do not do that work in this paper; rather, I merely show and explore some of the confusions and tensions and leave their possible resolution to legal scholars.

Dwelling on the second issue, that of white people’s potential responses to the judgment, may need more justification.⁶ Why, it may be asked, should we care about what white South Africans think about judgments that go against their

⁵ Throughout this paper, I use ‘black’ and ‘white’ in the inclusive sense familiar in South Africa. That is, they capture the crude and broad distinction that apartheid law made between whites and non-whites. Though there are important differences within the broad group ‘black’ (eg Indians, Coloureds) this dichotomy is still apposite in South Africa and clearly marks the different power relations in the country. This usage would be frowned on in, for instance, the United States, where differences between less privileged groups are stressed.

Whites are certainly not a monolithic group either, and there may be great differences in the way individuals respond to their place in South Africa. Furthermore, there may be differences between the white Afrikaans and white English communities, as well as between different socio-economic classes within the white population. I acknowledge this, and hope that my discussion in Section III brings out the necessary nuance.

For a study of perceptions of Employment Equity practices in South Africa, see RM Oosthuizen and V Naidoo ‘Attitudes Towards and Experience of Employment Equity’ (2010) 36 *South African Journal of Industrial Psychology* 836.

⁶ My focus on white South Africans continues the exploration of whiteness in my work over the past few years. For a defence of this focus see S Vice ‘How Do I Live in this Strange Place?’ (2010) 41 *Journal of Social Philosophy* 323; S Vice ‘Reflections on “How Do I Live in This Strange Place?”’ (2011) 30 *South African Journal of Philosophy* 503; S Vice ‘Race, Luck, and the Moral Emotions’ forthcoming in P Taylor, L Alcoff & L Anderson (eds) *The Routledge Companion to the Philosophy of Race*; and S Vice ‘Essentialising Rhetoric and Work on the Self’ (2016) 45 *Philosophical Papers* 103.

individual interests or seem disrespectful to their special value as human beings? They are certainly on no moral high ground from which to issue complaints and they need to come to terms with their status as a minority group with a morally tainted history. The Constitution promises them equal rights and protections as it does for any citizen, but it also requires that the historical and ongoing injustices and inequalities be rectified. Given that they were (and some would say, still are) the beneficiaries of that injustice, they cannot complain now if particular cases go against them. If the judgments are legally sound and if they are treated respectfully, there is nothing more to be said.

There are practical and political considerations that would make such a dismissive response naïve – for example, the fact that while whites may be stripped of political power, they still possess much of the wealth of the country and still therefore matter at least economically; or the more ethical and personal facts of their living in South Africa and thinking of it as home, and the importance for racial reconciliation of building a common sense of nationhood. In the current charged atmosphere of racial tensions and protests,⁷ understanding themselves and their position, at least, seems ethically and psychologically required. Furthermore, while I am certainly interested in whites' position in the country, as a white person myself, our position and difficulties are also a particular instance of general ethical issues in which I am interested. One is the conflict between respecting the dignity of unique, valuable persons, on the one hand, and on the other, realising structural reforms that would compensate a group of people for past injustices and ameliorate their current situation. A second is how all of us can and ought to think of ourselves as both unique individuals and members of socially and politically significant groups. A third issue, affecting *Barnard* more generally throughout independently of the racial dynamics, is the longstanding debate about whether there can be real, deep conflicts between values, and about how to adjudicate apparent conflicts. These fundamental ethical issues, which lie beneath and inform *Barnard*, make the case philosophically as well as legally interesting and explain some of the difficulties attending the judges' reasoning and verdict.

My plan is as follows: in the next part, I explore how the judges understand 'dignity', 'equality' and 'fairness' in their concurring judgments on *Barnard*, and how they understand the (ostensible) conflict amongst those values. I place their views in the contexts of longstanding philosophical debates over conflict between values, and the Kantian tradition that informs their interpretation of these values. Against this background, I then explore possible white responses to the judgment in *Barnard* in Part III. Throughout, my framework draws on the liberal tradition in political and moral philosophy, though I remain optimistic – without argument here – that many of that tradition's suggestions and prescriptions are not merely parochial. My explorations are philosophical rather than legal; I hope only to add a different perspective on the issues that many eminent legal scholars have already grappled with.

⁷ I have written this paper in the midst of ongoing student protests over the cost of higher education and its lack of transformation.

II VALUES AND CONFLICT

At issue in *Barnard* was whether the National Police Commissioner's decision not to promote Ms Barnard, a white woman, unfairly discriminated against her on the ground of race and thus contravened s 9 of the Constitution and s 6 of the Employment Equity Act.⁸ A member of the South African Police Services (SAPS) since 1989, Ms Barnard applied for a position in the SAPS National Evaluation Service, after the post was advertised in 2005. Despite being short-listed, interviewed and declared the best candidate, she was not appointed because the Divisional Commissioner felt that insufficient attention had been paid to racial diversity. The decision was made to re-open the process and Ms Barnard re-applied. Once again, she was declared by an interview panel to be the leading candidate and her appointment was recommended to the National Commissioner, who declined to appoint her on grounds, once again, that insufficient attention had been paid to racial representivity.⁹ The post remained unfilled in 2005, and was later withdrawn. Represented by the trade union Solidarity, Ms Barnard went to the Labour Court, and challenged the failure to promote her on grounds of unfair discrimination on the basis of race. The Labour court ruled in her favour.¹⁰ The SAPS appealed the ruling, and the Labour Court of Appeal upheld the appeal.¹¹ Ms Barnard then took the case to the Supreme Court of Appeal, which again ruled in her favour, set aside the Labour Court of Appeal ruling, and with some amendments, reinstated the original Labour Court ruling.¹² The SAPS appealed the decision to the Constitutional Court, which found in its favour, overturning the Supreme Court of Appeal's previous judgment. The main judgment by Moseneke ACJ, writing for the majority, concluded that the alleged discrimination was justified because the Police Commissioner's hiring decisions qualified as a restitutionary measure to further equality, as set out in s 9(2) of the Constitution, and that this complied with the requirements of the Employment Equity Act. The other justices agreed with the outcome, but some differed in their reasons, and in what they took the Act to require. Cameron and Froneman JJ and Majiedt AJ disagreed with Moseneke ACJ that the Act requires only that the measures be rationally related to their purpose; they argued that the Act requires, in addition, that such measures meet the standard of fairness. Van der Westhuizen J used a proportionality analysis to balance the competing interests in the case and to weigh up the importance of the affirmative action measures in this case against the impact on individual rights.

The judges realised the difficulties and importance of the case, especially regarding the need to balance the claims of equality and of dignity. In his main judgment, Moseneke ACJ writes that we are 'seized with a dispute over pressing

⁸ Employment Equity Act 55 of 1998.

⁹ The Constitutional Court only considered the second occasion.

¹⁰ *Solidarity obo Barnard and Another v South African Police Services* [2010] ZALC 10, 2010 (10) BCLR 1094 (LC) (Labour Court judgment).

¹¹ *South African Police Services v Solidarity obo Barnard* [2012] ZALAC 31, 2013 (3) BCLR 320 (LAC) (Labour Appeal Court judgment).

¹² *Solidarity obo Barnard v South African Police Service* [2013] ZASCA 177, 2014 (2) SA 1 (SCA) (Supreme Court of Appeal judgment) per Navsa ADP, with Ponnann JA, Tshiqi JA, Theron JA and Zondi AJA concurring.

concerns of equality and non-discrimination – matters of considerable personal and public importance’.¹³ Cameron and Froneman JJ and Majiedt AJ similarly consider the case important ‘since frank acknowledgment of these tensions is necessary to allow our society to move forward and to ensure a rational discussion that provides hope for the future for all’.¹⁴

I want now to set out the relevant values and issues with which the judges are concerned. Most generally, the issues at stake relate to the values of equality and dignity, and the tensions between them. Each individual person’s dignity must be respected, but at the same time equality must be progressively implemented. Sometimes, it appears that these two ends cannot be realised together, as furthering equality could in particular instances require treatment that, at least apparently, undermines the dignity of another party. A narrower tension is between the equality measures that are necessary to restore the dignity of those previously disadvantaged, and the dignity of those adversely affected by those measures. As dignity provides the justification for the discriminatory equality measures, this is a clash between ‘dignities’. A further two tensions, which I note but shall not discuss, are between service delivery and equality, and between service delivery and dignity. Delivering efficient service can pull against equality, if the person who can deliver is not one of the previously disadvantaged; and it can pull against dignity, if the person overlooked because of equality has his dignity undermined.

Making these tensions explicit reveals complexities in the moral issues at stake, disagreement about the meaning of the guiding values, and the inevitability of conflict between different values. I will be referring to most of the judgments in *Barnard*, without exploring the justices’ different arguments for their concurring judgments.¹⁵ I take liberally from the judgments, treating them simply as a text to be explored in itself, and making no pretence of adequately engaging with the details of each, the differing reasons the judges give for their shared judgment, and with the vast jurisprudential literature on these issues.

A Dignity and Equality

Famously, the new South Africa is founded on ‘[h]uman dignity, the achievement of equality and the advancement of human rights and freedoms’.¹⁶ These values, though not defined, are mentioned frequently, and they underlie and give substance to the rights in the Constitution.¹⁷ Section 9 of the Constitution is concerned with the equality right:

9. (1) Everyone is equal before the law and has the right to equal protection and benefit of the law.

¹³ *Barnard* (note 2 above) at para 4.

¹⁴ *Ibid* at para 74.

¹⁵ I do not refer to Jafta J’s concurring judgment, because it is more concerned with a technical legal issue (on cause of action) than with the substantive issues that interest me here.

¹⁶ Constitution, s 1.

¹⁷ Perhaps this lack of definition is appropriate and intentional, given the kind of document the Constitution is – one that needs to speak to many different parties, and which sought to bring to the table all the still conflicting parties.

(2) Equality includes the full and equal enjoyment of all rights and freedoms. To promote the achievement of equality, legislative and other measures designed to protect or advance persons, or categories of persons, disadvantaged by unfair discrimination may be taken.

(3) The state may not unfairly discriminate directly or indirectly against anyone on one or more grounds, including race, gender, sex [and many other grounds].

There are three points to note about how equality is used here. First, it is something to strive towards, something that our society has not yet reached. So it is a value that can be lost and regained, or that was never there to begin with – for example, the situation of non-whites during apartheid. Moseneke ACJ is careful to speak of the ‘achievement of equality’.¹⁸

Second, equality includes legal equality and equality of (legal) rights and freedoms, which were not available to blacks under apartheid. Presumably, however, these are founded on, and justified by, the *moral* equality of persons; this is not explicitly stated, but sometimes the language of equality in the judgments inclines more towards the moral than the socio-economic and legal sense.¹⁹ If people were not morally equal in the sense of being of equal importance from the moral point of view, and deserving equal consideration in all matters that concern them, the other kinds of equality would not be required. Moral equality cannot be lost, then, though it can be ignored.

Third, s 9(2) suggests a wide reading of equality: legislative and ‘other measures’ can be taken to ‘protect or advance’ the disadvantaged. As policies of the ANC government and *Barnard* suggest, the equal ‘rights and freedoms’ include economic and social equality, and measures ‘other than’ legislative ones can be taken to ensure this (though what exactly these other measures are is not stated). As Rósaan Krüger writes, the Constitution has a substantive view of equality, a view which ‘takes social and economic conditions of groups and individuals into consideration when determining the meaning of equal treatment’, and tries to undo long-standing patterns of disadvantage.²⁰

The Constitution is even briefer about the right to dignity. It states: ‘Everyone has inherent dignity and the right to have their dignity respected and protected.’

The value of dignity underlying this right is never defined, and the term has been understood in various ways by the Court.²¹ However, if we read the Constitution as informed by the long and familiar Liberal-Christian intellectual

¹⁸ *Barnard* (note 2 above) at para 28 (emphasis added).

¹⁹ See, eg, *Barnard* (note 2 above) at paras 429, 30, 176 and note 194.

²⁰ R Krüger ‘Equality and Unfair Discrimination: Refining the *Harksen* Test’ (2011) 128 *South African Law Journal* 479.

²¹ ‘Dignity’ is a complex term and can have other meanings in other contexts, as Christopher McCrudden’s useful history of the idea shows (C McCrudden ‘Human Dignity and Judicial Interpretation of Human Rights’ (2008) 19 *European Journal of International Law* 655). Stu Woolman’s thorough exploration of dignity suggests (controversially, I understand) that the notion has been understood in five main ways in the Court’s jurisprudence and has been used in three ways (S Woolman ‘Dignity’ in S Woolman & M Bishop (eds) *Constitutional Law of South Africa* (2nd Edition, OS, 2008) Chapter 36.2 to 36.3)) and has been used in three ways (Woolman, at 36.3). This variety, he argues, is organised by and draws on the central tenets of Kant’s ethics (S Woolman *The Selfless Constitution: Experimentalism and Flourishing as Foundations of South Africa’s Basic Law* (1st Edition, 2013)). The importance that the Court places on dignity is clear in *Dawood and Another; Shalabi and Another; Thomas and Another v Minister of Home Affairs and Others* [2000] ZACC 8, 2000 (3) SA 936 (CC) at paras 27–39.

tradition, and by Kant's view of human value, as the Court and scholars tend to do, 'inherent dignity' is a special worth that comes from and along with being a human being.²² 'Human' is a normative term – to be 'human' is already to be valuable, to have equal moral status, and is not a merely biological, non-normative term. Different accounts of dignity will ground this special value in different features. The Court has explicitly referred to and drawn from Kant, who grounds dignity in our rational nature.²³ 'Rationality' for Kant is a substantive, rather than merely instrumental ability; it is our ability to act for reasons, and to set ourselves ends through reason. Further, we are creatures who need not simply follow our natural instincts and desires (our 'empirical nature'), but who can rationally reflect on them and choose whether to indulge them. More generally, we can think of it as our ability to 'give an account of ourselves' to each other, as equal members of a moral community – an ideal Kant calls the 'kingdom of ends'.²⁴ Laurie Ackermann's expansive understanding of human dignity, draws on this substantial notion of rationality. He says that dignity

[arises] from all those aspects of the human personality that flow from human intellectual and moral capacity; which in turn separate humans from the impersonality of nature, enables them to exercise their own judgment, to have self-awareness and a sense of self-worth, to exercise self-determination, to shape themselves and nature, to develop their personalities and to strive for self-fulfilment in their lives.²⁵

Of particular importance to the Court is Kant's central tenet, that the dignity of individuals ought to be respected and that doing so rules out treating people as 'mere means' or 'tools' for furthering others' interests.²⁶ Judgments must therefore treat all relevant parties as 'ends in themselves' and not as *mere* means for the achievement of some social or personal good. How to distinguish between treating a person as a *mere* means, and treating her (sometimes permissibly) as a *means*, is important in *Barnard*, as we shall see.²⁷

²² Most notable in South Africa, is Laurie Ackermann's exploration of dignity and its intellectual traditions, which draws extensively on Kantian ethics: L Ackermann *Human Dignity: Lodestar for Equality in South Africa* (2012).

²³ See, eg, I Kant *Groundwork of the Metaphysics of Morals* (1785) (trans M Gregor, 1997) 4:428. This is the usual way of taking his claim, but sometimes Kant speaks as if it is autonomy (our ability to set ourselves the moral law) rather than rationality that is the ground of dignity; sometimes as if it is our moral dispositions; sometimes that only dutiful people (those with a 'good will') have dignity.

Rationality is not the only candidate for having dignity, and so, special value. One might instead think that being sentient, or the beloved creations of God, or merely being human, is the feature that confers moral status and special worth. Sentience, the feature chosen by the utilitarian moral tradition, has the benefit of including some animals, as well as human babies and human adults who are not rational; perhaps the theistic view is compatible with including animals. For an influential utilitarian approach to the moral status of animals, see work by Peter Singer, eg, his classic *Animal Liberation* (1975). For the importance of being human (in a rich, normative sense), see R Gaita *A Common Humanity* (2002).

²⁴ Kant (note 23 above) at 4:433f.

²⁵ Ackermann (note 22 above) at 23–24.

²⁶ Kant (note 23 above) at 4:428f.

²⁷ Kant allows that people may be treated as means; as social creatures who need each other's help to fulfil our ends, we do all the time. I treat the electrician as a means to have safe wiring; my students treat me as a means towards an education. This is permissible as long as we do not consider others significant only because they play these useful roles, and only if our interactions are also respectful of their dignity.

To be human, then, is to be valuable in a special way; one does not have to prove one's worthiness or earn dignity; it inheres in one regardless of what one does, solely in virtue of being rational. The Constitution and the Court's judgment in *Barnard* do not use this term, but perhaps they agree that dignity is also 'inalienable' – it cannot be lost or given up; there is nothing one can do that would remove one's dignity, and nothing others can do either. If this is so, then unlike legal equality and equality of (legal) rights and freedoms, but like moral equality, dignity can only be treated as if it did not exist, and that might have the effect of making people feel as if they have no dignity.²⁸ 'Destroying' dignity and moral equality is therefore, strictly speaking, impossible. Black people under apartheid were treated *as if* they had no dignity and little moral status; they were treated in ways that disrespected their dignity and moral equality, but they had them, nonetheless – otherwise there would be nothing to disrespect, and no harm in that attitude. The 'right to have their dignity respected', which the new Constitution promises them, would be unnecessary. It is thus interesting to note that the Constitution does not guarantee a 'right to dignity' but a 'right to have dignity protected and respected'. Apartheid was wrong (partly) because the value of black people was ignored and they were treated as moral inferiors. When we talk, as the judges do, of 'infringing', or 'undermining' or 'destroying' dignity, we mean attitudes and treatment that are not appropriate towards the special kind of value that humans have. So we can take appropriate or inappropriate attitudes towards a person, that express acknowledgement or lack of acknowledgement of the value that she nonetheless has, however she feels and however she is treated.²⁹ The judges are not always careful about this, and when I adopt their language and talk about 'denying', 'sacrificing' or 'invading' dignity, I am doing so loosely, and I mean behavior or attitudes that express disrespect for dignity and equality. For ease, I shall use 'disrespectful' or 'disrespect' to stand in for all such behavior and attitudes.³⁰

Our dignity, then, demands respectful responses and behavior, and places limits on what can be done to us. Our moral equality grounds legal and socio-economic equality, and the government must progressively realise the socio-economic rights the latter value justifies. While moral equality and dignity are properties of individuals, legal/socio-economic equality is most naturally (but not necessarily) a property of certain defined groups of people, which already raises the possibility that the dignity of one person and the realisation of equality for a

²⁸ If one wants the notion of 'human' or 'natural' rights to underpin legal rights, then those rights, too, could not be lost, though they might be ignored. On the relation between dignity and rights, and the sense in which dignity can be lost, see A Gewirth 'Dignity as the Basis of Rights' in MJ Meyer & WA Parent (eds) *The Constitution of Rights: Human Dignity and American Values* (1992). One usage in which dignity *can* be lost or acquired is when we say that a person has a 'dignified bearing', or 'holds on to her dignity', or, simply has dignity or is dignified. This demeanour might be lost or retained in certain situations, and retaining it may be praise-worthy – 'dignity under fire'. This sense of dignity is closely related to self-respect and integrity.

²⁹ Christopher McCrudden calls this the 'relational element' of this conception of dignity. McCrudden (note 21 above) at 28.

³⁰ I also use 'human being' and 'person' interchangeably, though in other contexts they can significantly come apart, and when I speak of 'respecting persons', this is short hand for 'respecting persons' dignity'.

group could be opposed. I want now to explore how the apparently competing demands of dignity and equality are dealt with by the justices in *Barnard*, and what they understood by a *fair* conclusion to such conflict. I do so, not to disagree with their verdict in *Barnard*, but to understand better what is at stake. I do not think that any of my tentative conclusions would undermine the legitimacy of the judges' decisions in these kinds of cases; in fact, they would probably unduly add complexity to an already complicated situation. Nevertheless, ethics should be prepared to deal with more complications than law. Insofar as this case can stand as a test of affirmative action measures, both black and white South Africans have a stake in it. I will also place the judges' deliberations in the context of debates in philosophy about how to understand conflict between values in general, and how to understand the dignity of persons. In this paper I can do no more than gesture at these debates; I do not aim to resolve any of them.

B Conflicts of Value

Isiah Berlin wrote:

If the ends of men are many, and not all of them are in principle compatible with each other, then the possibility of conflict - and of tragedy - can never wholly be eliminated from human life, either personal or social. The necessity of choosing between absolute claims is then an inescapable characteristic of the human condition.³¹

He thinks that value pluralism is a 'truer and more humane ideal' than the view that there is some way to reconcile all the varied 'ends of men', all the ideals that give substance and shape to a human life.

In the philosophical literature, there is much debate about whether values are at a fundamental level one or many; about how, precisely, to characterise conflict between values; and about whether Berlin is correct in thinking that a full reconciliation is impossible. These issues are an important contributor to the deep disagreement between consequentialist and non-consequentialist or deontological normative theories, and between pluralists and monists within both.³² The axiology of the most influential form of consequentialism, utilitarianism, is monist and welfarist: there is one fundamental value, welfare (in some version or

³¹ I Berlin 'Two Concepts of Liberty' in H Hardy (ed) *Liberty* (2002) 166, 214.

³² In standard consequentialist normative theory, the states of affairs brought about through actions (or omissions) are all that is relevant for assessing the rightness or wrongness of actions. In principle, a person could be sacrificed for the greater good (though consequentialists have sophisticated ways of avoiding this in fact). Non-consequentialists or deontologists can admit that consequences are sometimes morally relevant, and that sometimes sacrificing one person for an important social goal might be justified (though Kant and some Christian philosophers would not admit this). However, they do not see consequences as the source of moral value and moral status, and sometimes such sacrifices will be impermissible, regardless of the good consequences.

another), and it ought to be promoted or maximised.³³ The axiology of Kantian ethics, the kind of non-consequentialist or deontological theory I focus on here, is also monist; as we have seen, the fundamental value is the dignity of rational nature, which must be respected in all our actions.³⁴ Both these theories make reference to their grounding value in situations of conflict, but partly because of the different responses they prescribe, these values operate in very different ways. I shall return to this point in relation to the conflicts in *Barnard*.

Bernard Williams agrees with Berlin that value-conflict is ‘necessarily involved in human values, and [is] to be taken as central by an adequate understanding of them’, but he also thinks that it is by no means clear what it is ‘for values to be plural, conflicting and irreducible’.³⁵ Putting aside internal or logical inconsistencies between values and concentrating on contingent conflicts, there are at least two ways of thinking about them. First, one could think that values are not comparable or commensurable,³⁶ in the sense that there is no higher-order or ‘super-value’ by which conflicts between lower-order values could be settled. A utilitarian would of course deny this: ‘welfare’ in whatever sense of the term settles conflict, so strictly speaking, there *is* no conflict fundamentally. When there is a conflict in a particular situation between, say, the demands of justice and the demands of love, we settle it by calculating which choice would maximise welfare. Conflicts are only ever apparent and there will be a correct answer about how to settle them. If your choice maximises welfare, then it is required; if another available choice would produce more welfare, it is impermissible not to choose it; if more than one choice would produce equal welfare, either is unproblematically permissible. There is no reason to feel guilt or regret over the option not taken.

A deontologist can agree that there might be one correct answer or less strongly, an all-things-considered right answer to a conflict. Kant, for instance, would

³³ Utilitarianism understands ‘utility’ or ‘the social good’ in terms of happiness, desire satisfaction, or welfare. JS Mill’s *Utilitarianism* is the classic text (many editions; available at <http://www.gutenberg.org/files/11224/11224-h/11224-h.htm>). There can, however, be non-utilitarian consequentialist theories (eg GE Moore’s theory in GE Moore *Principia Ethica* (Revised 1st Edition, 1929) and GE Moore *Ethics* (2nd Edition, 1966), unhelpfully called ‘ideal utilitarianism’ – unhelpful because it is not, strictly speaking, utilitarianism).

In this paper, I consider only standard or classical consequentialism, and ignore later developments which depart from the strictly maximising approach to value, and which are sensitive to factors like rights, distribution, and the nature of the actions themselves. I do so because the conflict facing the judges in *Barnard* seems to be a fairly straightforward conflict between the deontological value of dignity and the urgent demand of realising social and economic goals to improve the welfare of the majority of South Africans. I am not sure that introducing complications to consequentialism here would help. For developments in consequentialism, see M Slote ‘Satisficing Consequentialism’ (1984) 58 *Proceedings of the Aristotelian Society* 139; A Sen ‘Consequential Evaluation and Practical Reason’ (2000) 96 *The Journal of Philosophy* 96, 477; and ‘Rights and Agency’ (1982) 11 *Philosophy and Public Affairs* 3.

³⁴ I focus on Kant’s ethics in this paper, but there are other examples of non-consequentialist moral theories – eg WD Ross’s in WD Ross *The Right and the Good* (1930), or TM Scanlon’s contractualism in, for instance, TM Scanlon *What We Owe to Each Other* (1998).

³⁵ B Williams ‘Conflicts of Values’ in *Moral Luck* (1981) 71, 72–73. See also B Williams ‘Ethical Consistency’ (1965) 39 *Proceedings of the Aristotelian Society, Supplementary Volumes* 103.

³⁶ While some philosophers keep these distinct, I will use the terms interchangeably. See J Raz *The Morality of Freedom* (1986) and R Chang ‘Introduction’ in R Chang (ed) *Incommensurability, Incomparability, and Practical Reason* (1997).

require strict compliance with the negative duties of right, which are grounded in respectful treatment of rational creatures; reason will arrive at an answer of how to act, and cannot, without internal incoherence, arrive at competing answers.³⁷ Other deontologists may not explain resolutions in terms of some 'super-value', or think that conflict would dissolve on rational reflection. Perhaps, in a particular situation, there is an all-things-considered best option, and all would agree that one did the right thing. However, this does not mean there is in fact only one thing that is right to do, nor that there is another more general value that settles the conflict, nor that the normative force of the rejected option evaporates. There can be more than one genuine obligation, and the moral force of the options not taken remains. In genuinely tragic situations, it can be the case that whatever a person does will be wrong, and there is no clear all-things-considered right choice. In these non-consequentialist interpretations of conflict, some 'moral residue' will be left, some sacrifice incurred. Regret may be a rational response, even though one has chosen in the best way possible.³⁸

Second, there is another way of interpreting incommensurability that is available to non-consequentialists: rather than concerning whether there is a common measure or super-value by which to adjudicate conflict (a vertical model), it concerns whether two values can meaningfully be compared with each other (eg dignity and equality), or two instances of one value (the dignity of Ms Barnard and the dignity of those harmed by apartheid) – a horizontal comparison. To compare the dignity of one against the dignity of another, or many others, or the dignity of one against the inequality of a group, would be like trying to compare apples and oranges. More strongly, in ethically difficult cases it can sometimes be morally dubious or distasteful even to try. For non-consequentialists, if we do have to make a choice in these situations, the value of the option that was not chosen again remains and may exert normative force over us still; again, there will be moral residue and so grounds for regret. Consequentialists would reject incomparability in this sense and claim that welfare is always available to settle the conflict.

Against this background, let us return to *Barnard*. The Constitution states that no rights are absolute and that they may need to be limited if that can be shown to be 'reasonable and justifiable in an open and democratic society'.³⁹ However, the limitations must be based on certain kinds of reasons, which themselves need

³⁷ Kant does not think there can be genuine conflicts between the necessary, *a priori* duties generated by the categorical imperative. On this, see J Timmermann 'Kantian Dilemmas? Moral Conflict in Kant's Ethical Theory' (2013) 95 *Archiv für Geschichte der Philosophie* 36. The contrast between 'duties of right' and 'duties of virtue' is from Kant *Metaphysics of Morals* (1797) (trans M Gregor 1996). When duties of right clash with duties of virtue, the former should be followed.

³⁸ See Williams 'Ethical Consistency' (note 35 above). On tragic conflicts, see RB Marcus 'Moral Dilemmas and Consistency' (1987) *The Journal of Philosophy* 121; H Richardson *Practical Reasoning About Final Ends* (1994); M Nussbaum 'The Costs of Tragedy: Some Moral Limits of Cost-Benefit Analysis' (2000) *Journal of Legal Studies* 1005.

³⁹ Constitution, s 36.

to take into account a number of factors beyond outcomes.⁴⁰ The Constitutional Court also makes use of a proportionality test to resolve conflict, which may or may not include the balancing of conflicting interests and rights.⁴¹ The rights to dignity and equality can therefore in principle be limited or balanced against each other, so the fact that at least one may need to be infringed in a situation of conflict is not in itself surprising – though some might disagree that certain of those rights should ever be infringed. However, because of the way that equality and dignity are understood by the Court, and because of the facts of post-apartheid South Africa, it is almost inevitable that they will conflict, and resolving the conflict in a ‘reasonable and justifiable’ way with no moral residue will be, as I try to show, very difficult.

I begin with dignity. That the Court relies on a Kantian framework to give content to dignity introduces a difficulty and a tension that is not resolved in *Barnard*, and could not be. There are in fact two claims that are considered in the judgments. The first is that one person’s dignity cannot be weighed against another person’s, nor many others. We see this when Van der Westhuizen J calls the ‘balancing’ of one dignity against another inappropriate. Ms Barnard’s dignity cannot be ‘weighed’ against that of others, and neither can the dignity of the millions of victims of apartheid be weighed against her dignity.⁴² The second is that dignity cannot be weighed against another value, like equality. Both claims seem to be understood through a Kantian lens.

On the first claim: Kant distinguishes dignity, which has worth, from everything else, which has a price.⁴³ This means ‘at least that whenever one must choose between something with dignity and something with mere price one should always choose the former’.⁴⁴ Things with price can be traded for each other and measured against each other; dignity cannot, and must always take precedence. This makes his ethics deeply antithetical towards consequentialist reasoning. One reason Kant thinks this is that rational nature, which grounds our dignity, is also the ground and condition of *all* value; there would be no value in the world without rational nature, which confers value through its choices.⁴⁵ While it is clear, then, that dignity takes precedence over things with price, it is not so clear whether Kant thought dignity could be measured against dignity. As we saw above, however, at least one Constitutional Court judge denies that this is appropriate. Weighing the value of each person against each other or against

⁴⁰ Section 36(1) lists the following factors as relevant to considering whether a rights limitation would be justifiable: (a) the nature of the right; (b) the importance of the purpose of the limitation; (c) the nature and extent of the limitation; (d) the relation between the limitation and its purpose; and (e) less restrictive means to achieve the purpose.

⁴¹ See N Petersen ‘Proportionality and the Incommensurability Challenge – Some Lessons from the South African Constitutional Court’ (2013) *New York University Public Law and Legal Theory Working Papers* Paper 384. Van der Westhuizen J uses a proportionality analysis in his judgment in *Barnard* (note 2 above).

⁴² *Barnard* (note 2 above) at para 178.

⁴³ See Kant (note 23 above) at 4:434–5 (‘In the kingdom of ends everything has either a *price* or a *dignity*. What has a price can be replaced by something else as its *equivalent*; what on the other hand is raised above all price and therefore admits of no equivalent has a dignity.’)

⁴⁴ TE Hill *Dignity and Practical Reason in Kant’s Moral Theory* (1992) 48.

⁴⁵ This is Christine Korsgaard’s interpretation of Kant (note 23 above) 4:428–9. See C Korsgaard ‘Aristotle and Kant on the Source of Value’ (1996) *Creating the Kingdom of Ends* 225, 240–241.

another value is impermissible and, strictly speaking, impossible. Dignity is not the kind of value individual instances of which can be weighed or traded against each other, and there is no common, higher-order measure against which trade-offs between it and another value could be made.

On this interpretation of the special value of dignity, comparing or weighing up the dignity of Ms Barnard and the dignity of the many people disadvantaged by apartheid is confused, and is to misunderstand what dignity is. That is, you can do it, and the justices do, and you could come to an all-things-considered conclusion on that basis, but your conclusion will not be morally valid, because you have not been debating *dignity* at all. Further, these attempts are themselves expressions of disrespect, as you are not responding appropriately to a special kind of value.

The difficulties increase when the demands of socio-economic equality enter the picture, and this leads us to the second claim. Because the judges are at least in principle committed to a Kantian understanding of dignity, they face the challenge of responding appropriately, both to dignity and equality, and to their conflict. *Moral* equality is subsumed into dignity for Kant, but legal and socio-economic equality are distinct and subordinate. In principle, then, a strict Kantian would face no tension at all. No-one's dignity can be traded for any amount of socio-economic equality. Other deontologists need not follow Kant on this point, and I am not insisting that the judges should either, but insofar as they are assuming a Kantian framework in some sense, it is at least a question they need to consider. In fact, the judges depart from Kant in seeing the case as one of at least apparent conflict about which they must reach a verdict, if not a solution. They retain the spirit of Kantian ethics, however, in not wanting to resort to crude consequentialist calculations that might without fuss or moral residue sacrifice an individual person's dignity for the goal of equality.⁴⁶ However, their judgments sometimes present the resolution of the conflict as if it would incur *no* loss – the decision against Ms Barnard is not unfair after all, and so is not disrespectful – and sometimes as if it would involve a sacrifice but one that is all-things-considered justified because of the importance of the other goal of equality. Cameron and Froneman JJ and Majiedt AJ, for instance, recognise the 'perils' of remedial action, which 'may exact a cost our racial history demands we recognise'.⁴⁷ However, they nonetheless conclude that the Police Commissioner's decision not to appoint Ms Barnard was fair, and that 'it is not necessarily an injury to dignity to view a person only through the lens of one ground listed in section 9(3), provided the reason for doing so is to redress historical inequality'.⁴⁸ Van der Westhuizen J admits the possibility of loss, and the possibility that the rights to and values of equality and dignity might sometimes compete. He writes:

[a]spects of a person's right to dignity may sometimes have to yield to the importance of promoting the full equality our Constitution envisages. Other times, the impact of

⁴⁶ How to give due consideration to consequences in a principled way is a long-standing problem for non-consequentialists. See, eg, P Foot 'The Problem of Abortion and the Doctrine of the Double Effect' (1978) *Virtues and Vices and Other Essays in Moral Philosophy*.

⁴⁷ *Barnard* (note 2 above) at para 79; and see para 93.

⁴⁸ *Ibid* at para 117; and for the judgement that the decision was fair see 121, 123.

equality-driven measures with laudable aims may not be justifiable in view of severe damage to human dignity.⁴⁹

Further complicating matters is that in *Barnard* equality is sometimes understood as a condition for dignity; that is, equality is important because inequality is disrespectful to dignity. As Van der Westhuizen J says, '[m]easures to achieve equality are supposed to restore dignity'. However, he continues, 'their practical implementation could also impact on the human dignity of individuals'.⁵⁰ Here, presumably, 'equality' means legal and socio-economic equality. As this suggests, the relation between equality and dignity is not at all clear, neither in *Barnard* nor in other Court jurisprudence.⁵¹ However, one relation here, it seems, is that the lack of equality leads to a diminished sense of dignity, and so working towards equality is one way of respecting dignity.⁵² On the other hand, working towards equality in order to revive a sense of dignity can also lead to those who are negatively affected having *their* sense of dignity diminished.⁵³ And then one has to weigh the dignity of one with the dignity of others, in the name of both dignity and equality. Confusingly, Van der Westhuizen J denies that one person's dignity can be 'balanced' against another,⁵⁴ but right afterwards says that the calculation required to restore the dignity of some (via equality measures) at the cost of the dignity of others was done when the Constitution was agreed on.⁵⁵ Setting aside the apparent contradiction here (between not balancing, on the one hand, and calculating, on the other), this must mean that the Constitution from the start allows us to undermine the dignity of some if it is a (presumably necessary, indispensable and proportionate) means towards restoring the equality and thus the dignity of many others who were historically disadvantaged – in that situation dignity would not be unduly or unfairly disregarded. It would not be possible to realise equality if such a trade-off were ruled impermissible in principle.

That socio-economic equality is understood substantively and as something to be progressively realised, further contributes to the conflict of values the judges have to deal with. It requires them to take future-directed considerations into account –

⁴⁹ Ibid at para 169.

⁵⁰ Ibid at para 178; and see para 176.

⁵¹ The connection between equality and dignity is not clarified in the Constitution, and there is ongoing debate about which, if either, is the more fundamental value, and whether one grounds the other. In other judgments, the Court has taken 'equality of dignity' to be basic. Rósaan Krüger writes: 'In the few years of constitutional democracy preceding the enactment of the Equality Act, the equality jurisprudence of the Constitutional Court of South Africa ... firmly established human dignity as the interest protected by the equality right, and therefore as the interest at the core of the prohibition of unfair discrimination' (R Krüger 'Small Steps to Equal Dignity: The Work of the South African Equality Courts' (2011) 7 *Equal Rights Review* 27). Justice Laurie Ackermann makes dignity the fundamental value and guiding ideal of the Constitution in Ackermann (note 22 above). For a helpful account of Ackermann's work on the connection, see C McConnachie 'Human Dignity, "Unfair Discrimination" and Guidance' (2014) 34 *Oxford Journal of Legal Studies* 609. McConnachie is sceptical that dignity can play the role that Ackermann wants.

⁵² Whether equality, then, would be subordinate to dignity is not clear.

⁵³ My understanding of Van der Westhuizen J's claim in this paragraph does not support understanding equality in terms of equal dignity, ie we cannot say: 'What are we supposed to be equal in respect of?' and answer: 'Dignity', if equality can both undermine and enhance dignity.

⁵⁴ *Barnard* (note 2 above) at para 178.

⁵⁵ Ibid.

the ongoing requirement to realise equality – while at the same time responding to a present, already existing, value which places categorical constraints on decisions – the constraint imposed by the dignity of actual persons as ends-in-themselves. Rather than bring about a state of affairs which might not otherwise have existed, Kantian (and other deontological) ethical theories require us to respond with respect to the valuable creatures already existing. Their existence prevents us from bringing about some of the states of affairs consequentialist calculations would demand. The justices are required by the Constitution to permit measures (meeting certain conditions) that will realise substantive equality in the future, but doing that might require the kind of interference or disrespectful attitudes that are contrary towards a (some would say *the*⁵⁶) foundational, deontological, value of the Constitution. As suggested earlier, the Court presumably cannot interpret the Constitution such that any conflict between equality and dignity must come down on the side of dignity, for that would be to render impossible the Constitution's transformational and egalitarian goals. This means that the Court is from the start open to, and indeed required to make, consequentialist calculations, despite its also being required to respect deontological values.

I have stressed that on the Kantian understanding of dignity, which infuses the judgments in *Barnard*, dignity is not comparable; one person's dignity cannot be weighed against another's or many others', nor against another value like equality. And I have stressed the complications that arise when the ideal of equality, which requires future-directed action, conflicts with the dignity claims of actual, existing ends-in-themselves. I have made much of these points, because they show the kinds of axiological complexities and sacrifices that are inevitable in cases like these.

We have (at least) three options in the face of this: First, 'dignity' could be given a different, non-Kantian interpretation (one that makes it a properly comparable notion). The work necessary for this still needs to be done, although ongoing work on the notion of Ubuntu might yield results.⁵⁷ Second, we could give up on all attempts to 'weigh' or 'balance' or 'trade' dignity with another value. Dignity 'trumps' all other considerations.⁵⁸ What to do when one person's dignity comes into conflict with another's, or that of many others, is not clear; perhaps this is where numbers legitimately count. However, it seems difficult to see how courts or government bodies could practically do without comparisons of value. Conflicting needs and conflicting rights, scarcity of resources, political impossibilities – all these mean that hard choices must be made. For those of a sturdy consequentialist temperament, the choices might be made with no sense of regret or sacrifice, but such purity seems incredible in the face of the wrenching decisions the judges are required to make. It is difficult to see how a

⁵⁶ See note 51 above.

⁵⁷ See D Cornell 'Is There a Difference That Makes a Difference Between uBuntu and Dignity?' (2010) 25 *South African Journal of Public Law* 382; Y Mokgoro & S Woolman 'Where Dignity Ends and uBuntu Begins: An Amplification Of, as well as an Identification Of a Tension in, Drucilla Cornell's Thoughts' (2010) 25 *South African Journal of Public Law* 400.

⁵⁸ See R Dworkin 'Rights As Trumps' in J Waldron (ed) *Theories of Rights* (1984) 153; and Robert Nozick on 'side constraints' in R Nozick *Anarchy, State, and Utopia* (1974) 26 (the section titled 'Moral Constraints and the State').

theoretical standpoint that denies the possibility of deep conflict, rational regret and unfortunate sacrifice would have any practical force; and it is difficult to see that it captures the phenomenon of a morally and politically complex country like South Africa. Third, one might accept that one is doing something inappropriate when one tries to weigh one dignity against another value or another dignity and accept it as the price we pay for a history of inequality; we have to work with a watered-down version if we are to work with it at all, and we have to work with it because the Constitution says so, as Moseneke ACJ trenchantly observes in another context.⁵⁹ Here, unlike the second option, we can see the situation in *Barnard* as a moral dilemma: whatever decision is reached, something of value will be sacrificed. There is no choice without moral residue. Recognising this would not help when making hard choices, of course, but it would guide the kinds of reasons one gives to justify one's choice, which as I explore below, is, on one understanding, important for the fairness of a decision. Admitting the sacrifice to those adversely affected would be a way of acknowledging that they matter and that they are losing something they reasonably want to retain.

While the judges grapple admirably with these issues, at the end of the day they resort to weighing and comparing considerations, even if they try not to be crude, and it is difficult to think of a more justifiable alternative. In a situation like ours in post-apartheid South Africa, dignity must be quantified and weighed against equality, and sometimes the dignity of one person must be weighed against the dignity of those people disadvantaged by apartheid. One conclusion that therefore suggests itself so far is that in the aftermath of systematic and long-lasting injustice, we may be required to act as if values that are strictly incomparable can be weighed or compared against each other; we may be required for political and practical reasons to think consequentially about a non-consequentialist value; or, we may have to disregard the dignity of one in the name of restoring and protecting the dignity and moral equality of many others. In the aftermath of injustice, we cannot always be just, and that is part of the injustice and its legacy.

This will strike many as too quick a conclusion. Surely, as the Constitution says and the judges emphasise, there are fair and unfair decisions in these cases. As long as the decision of the judges is fair, then the dignity of Ms Barnard is respected. A fair decision is not necessarily one all parties will agree with, but it is one in which there is no egregious sacrifice. The judgments in *Barnard* have a lot to say about fairness, and I end this section by looking briefly at this.

C Fairness

Dignity and equality are similar in the sense that they can both rule out or require discriminatory measures. That a discriminatory measure would be disrespectful towards dignity can rule it out; that it would respect dignity can require it, or provide overriding reasons for it. Similarly, that a measure would enhance or set back equality are reasons for and against it. However, discrimination on the basis of certain grounds must be justified – it must be shown to be *fair*. On one, strong

⁵⁹ *Barnard* (note 2 above) at para 37. He says this after observing that a measure that passes the three-fold *Harksen* test for fairness is fair, 'because the Constitution says so'.

interpretation of the Constitution (supported by s 9(5)), there is a presumption that discrimination is unfair unless shown to be otherwise; the burden of proof for its fairness is on those who wish to discriminate.⁶⁰ As I mentioned earlier, however, this would render the realisation of equality difficult. Matters are complicated because, as Cameron and Froneman JJ and Majiedt AJ note, some interpret the Constitution as making fairness itself an independent value, along with equality and dignity.⁶¹ This leads them to choose fairness as the appropriate standard to use, one that is importantly consistent with the aims of the Employment Equity Act, 'namely, to avoid over-rigid implementation, to balance the interests of the various designated groups, and to respect the dignity of rejected applicants'.⁶² I set this aside, however, as much of the judges' discussions concern whether the discrimination against Ms Barnard counted as fair or unfair, rather than whether fairness is a core constitutional value in its own right.

The concurring judgment of Cameron and Froneman JJ and Majiedt AJ makes fairness the appropriate standard of whether discrimination is justified or not.⁶³ Fairness seems to come down to whether the decision not to appoint Ms Barnard unduly impacted her dignity, and in order to avoid a vicious circle, 'unduly' cannot mean 'unfairly' or 'unjustifiably'. So we then have to investigate what 'undue' infringements of dignity are. The judges of course refer to other cases in which fairness was at issue, and in which tests for fairness were offered – for example, *Harksen* and *Van Heerden* – though they do not go into them in detail.⁶⁴ They refer to them, but their statements also make or suggest more fundamental points about fairness itself, and it is these, rather than legal tests for fairness, in which I am interested.

The past discrimination under apartheid was clearly unfair, but the Constitution allows that discrimination in post-apartheid South Africa could be fair. The state must realise equality, and it may do so by taking special measures to 'protect or advance persons, or categories of persons, disadvantaged by unfair discrimination'.⁶⁵ Special measures taken for this reason presumably might discriminate against other persons or groups of persons not previously disadvantaged, but that discrimination would not be unfair.

The main concern of the judges is that unfair discrimination may involve unjustified disrespect for dignity. However 'fairness' is tested, a fair verdict will express the appropriate attitudes and display appropriate treatment towards human beings. For example, Moseneke ACJ writes that measures that are 'directed at remedying past discrimination must be formulated with due care not

⁶⁰ 'Discrimination' is therefore used neutrally; an instance of discrimination can be fair or unfair. The Constitution is careful, in s 9(3)–(5), to prohibit *unfair* discrimination.

⁶¹ In para 98 they quote O'Regan J in *Mphahuli*: 'Fairness is one of the core values of our constitutional order' (*Lufuno Mphahuli & Associates (Pty) Ltd v Andrews and Another* [2009] ZACC 6, 2009 (4) SA 529 (CC), 2009 (6) BCLR 527 (CC)).

⁶² *Barnard* (note 2 above) at para 97.

⁶³ *Ibid* at para 98.

⁶⁴ *Harksen v Lane NO and Others* [1997] ZACC 12, 1998 (1) SA 300 (CC), 1997 (11) BCLR 1489 (CC) and *Minister of Finance and Another v Van Heerden* [2004] ZACC 3, 2004 (6) SA 121 (CC), 2004 (11) BCLR 1125 (CC). See Krüger (note 20 above).

⁶⁵ *Barnard* (note 2 above) at para 29.

to invade unduly the dignity of all concerned'.⁶⁶ The Employment Equity Act, from which the judges draw, says that employers must diversify the workforce 'based on equal dignity and respect of all people'.⁶⁷ This might explain the importance Cameron and Froneman JJ, Majiedt AJ, and Van der Westhuizen J, place on the reasons given for their judgment. They must be capable of being understood by those they affect, because '[k]nowing why the decision was adverse enables the aggrieved person to understand – an understanding that encourages participation in rebuilding our divided country'.⁶⁸ Though they do not state it in this way, we could say on their behalf that reasons must be 'public'. Public reasons make reference only to reasons that are in principle acceptable to all citizens; they do not bypass citizens' rational capacities through manipulation, lying or other distortion, and they do not depend on highly controversial assumptions.⁶⁹ Not providing reasons, or expecting a person to accept idiosyncratic or contested views as reasons for an unfavourable judgment against her, would take away from the legitimacy of the judgment and suggest that the law does not speak for or on behalf of all citizens. It would, in Andrew Lister's words, 'fail to respect... fellow citizens as persons, which is to say as beings capable of recognising and responding to justificatory reasons'.⁷⁰ The provision of public reasons is especially important in such a diverse country as South Africa, where citizens who have to live together have deep disagreements about religion, the good life, liberal and traditional values, and many other matters. Critics of public reason would say that for this same reason, it is extremely difficult to find such uncontroversial reasons.

The publicity of the reasons may therefore provide a threshold for fair discrimination; it would show that a person's dignity was not unduly undermined and so no insupportable sacrifice of values results from the decision. Other possibilities for such a threshold can be gleaned from the judges' reasoning. First, after saying that equality measures must not unduly invade dignity, Moseneke ACJ says that we must take care that the measures 'are not an end in themselves'; they are 'not meant to be punitive or retaliatory', but rather urge us towards a more equal and fair society.⁷¹ Second, Cameron and Froneman JJ and Majiedt

⁶⁶ Ibid at para 30; and see paras 31, 32.

⁶⁷ Ibid at para 42 (quoting s 15 of the Employment Equity Act). Jafta J argues that in fact, the Police Commissioner's decision can never have been considered unfair, because it follows s 6(2)(a) of the Employment Equity Act, which says that 'it is not unfair discrimination to take affirmative action measures consistent with the purpose of this Act' (ibid at para 208). This makes any discrimination consistent with the goal of 'achieving equity in the workplace' (Employment Equity Act s 2) already fair. If this is so, then the other judges' attempts to assess the fairness of the decision are unnecessary. Moseneke ACJ seems to make a similar point in paras 36–37, but in relation to the Constitution, rather than the Employment Equity Act. The judges also disagree over whether the fairness of the Employment Equity Plan itself is at issue, or the way that the Police Commissioner implemented it. I leave this point to legal experts, and for the sake of my ethical interests in this paper, do take into account what is said about fairness in the other judgments.

⁶⁸ Ibid at para 106, and see 111. Van der Westhuizen J makes a similar point at para 193.

⁶⁹ The notion of public reasons is especially associated with the work of John Rawls, eg in J Rawls *Political Liberalism* (Expanded Edition, 2005); J Habermas *Moral Consciousness and Communicative Action* (1983) (trans C Lenhardt & SW Nicholsen, 1999); and G Gaus *The Order of Public Reason* (2011).

⁷⁰ A Lister 'Public Reason and Moral Compromise' (2007) 37 *Canadian Journal of Philosophy* 1, 7. And see C Larmore 'The Moral Basis of Political Liberalism' (1999) 96 *The Journal of Philosophy* 599.

⁷¹ Barnard (note 2 above) at para 30.

AJ write that the Constitution allows remedial measures ‘because it recognises that substantive equality can be achieved only by providing advantages to groups of people upon whom apartheid imposed heavy disadvantages’.⁷² However, the motivation for these measures, while using the same racial classifications as apartheid, is ‘the opposite of what inspired apartheid: for their ultimate goal is to allow everyone to overcome the old divisions and subordinations’.⁷³ ‘It is not necessarily an injury to dignity to view a person only through the lens of one ground listed in section 9(3),’ they write, ‘provided the reason for doing so is to redress historical inequality’.⁷⁴ Third, Van der Westhuizen J asks whether the impact on Ms Barnard’s dignity was ‘reasonable and justifiable in light of the goal of substantive equality’,⁷⁵ and he answers that two factors need consideration in answering this: Was she treated as a mere means towards achieving the goal of equality? And was the measure taken by the National Commissioner an ‘absolute barrier’ to her career advancement?

We have, then, some possibilities for threshold criteria for fair treatment that do not unduly diminish dignity. Fair treatment is treatment that can be justified to the persons affected, in terms they can accept; discriminatory measures must not be chosen for their own sake but for the end of equality, must contribute to the goal of equality and reconciliation, and be motivated by that goal; the affected person must not be treated as a mere means towards the goal; and the treatment should not be an absolute barrier to a person’s career.

Whether any of these criteria are met might be difficult to discover in practice. One reason (though perhaps this is unfair) is that it is difficult to know the intentions and end of a decision and action. Did the National Commissioner *really* intend to contribute to the goal of equality? Identifying intentions is a murky business at best, but perhaps we can set this aside, or resort to an argument to the best explanation. Another reason is that it is difficult to judge in advance whether the decision in a particular case would in fact contribute to equality. This is an empirical matter to be settled in the future, once the effects of the decision are known. Whether there is an absolute barrier to a person’s career is also an empirical matter which may not be clear at the time.

These claims might be thought to be too quick: surely we can draw from experience and make reasonable predictions? Certainly, but at least in respect of legal decisions, we need to know in advance of waiting for the effects to play out whether our decision to allow them to play out was justified. Legal justifications cannot be hostage to fortune. It might also be said that I am ignoring the symbolic or expressive value of the judges’ decision. Their decision expresses, or is a symbol of, a commitment to equality, and so has great political significance in South Africa. Evidence for this interpretation is Van der Westhuizen J’s remark that even the *perception* that a person is treated as a mere means would undermine the pursuit of equality.⁷⁶ Presumably this is not (or not only) an empirical matter,

⁷² Ibid at para 93.

⁷³ Ibid.

⁷⁴ Ibid at para 117.

⁷⁵ Ibid at para 180.

⁷⁶ Ibid.

but a matter of what a commitment to ‘progressive realisation of equality’ consists in and requires. This is plausible, but the point pulls both ways – we can worry about the perceptions of both parties.

Finally, and perhaps most difficult to resolve given the explicitly acknowledged tension between equality and dignity, it is difficult to know when a person is being treated as a ‘mere means’. Ms Barnard is in some sense being evaluated as a means or an obstacle to furthering equality, otherwise, as all admit, she would have been appointed to the disputed position. But is she being treated as a *mere* means? The judges sometimes talk as if there is no real, or no significant, sacrifice of value if she is not *unduly* affected, but that does not help us here if by ‘unduly’ is meant ‘unfairly’ – we would be going in circles. We could also say on their behalf that she is not being treated as a mere means because her qualifications and her claim are given their serious attention. If she were being treated as a mere means, she need not even have been considered; in consequentialist reckoning, the plight of the disadvantaged in South Africa clearly outweighs, in brute numbers, the claims of one woman. That Ms Barnard’s claim is taken seriously by the Court shows its commitment to the deontological values in the Constitution, and a commitment against giving consequentialist reasoning automatic supremacy. The publicity criterion may be related to this: a person is not treated as a mere means when public reasons are provided; her rationality is not bypassed, but respected, even if she remains unhappy by the verdict. In explaining what Kant means by treating people as a mere means, Thomas Hill says that insofar as people are used as a means, ‘they must *be able* to adopt the agent’s end, under some appropriate description, without irrational conflict of will’.⁷⁷ This does not mean that they in fact *will* adopt the reasons; in particular situations that may be recalcitrant, irrational, overcome by resentment or fear of personal loss.

If we are seeking the criteria for acceptable compromises of dignity, therefore, the judgments in *Barnard* are inconclusive. There does not seem to be a clear, explicit and agreed upon interpretation of when discrimination is fair and when an injury to dignity is reasonably outweighed by claims of equality (assuming the two values can be weighed against each other). The references to *Harksen* and *Van Heerden* do not resolve the deeper philosophical uncertainties (and perhaps it is unfair to expect this from the judges). The publicity criterion provides a promising way to interpret what it means to treat a person as a mere means, but that criterion is nowhere explicitly stated or acknowledged, let alone defended, in *Barnard*. Appeals to fairness do not resolve what is at stake in *Barnard*.

III WHITE SOUTH AFRICA

In the rest of this paper, I further explore the tensions between equality and dignity in South Africa, focussing now on white South Africans and their reactions to discrimination cases like *Barnard*. I am particularly interested in how those negatively affected – people like Ms Barnard – would take judgments against them, and what is at stake politically and morally in such debates. Would they accept the Court’s verdict as fair (setting aside whether it *is* fair), and what

⁷⁷ Hill (note 44 above) at 45.

needs to be the case about them in order for that to be possible? Examining these questions takes us to issues of identity, and to how group membership may partly constitute identity.

A Claims of Identity

Dignity is in the first instance an attribute of individual human beings, and to respect 'dignity' is to respect each person, in virtue of a property he shares with every other person. In this sense, dignity is both universal and individual: we respect individual persons, in virtue of a universally shared property. However, under the influence of what has come to be called 'identity politics' or 'the politics of difference', this sense of dignity and respect has been extended.⁷⁸ People must now be respected not only in virtue of their common humanity, but also in virtue of their differences from one another, and in virtue of the properties they share with only certain others. In this context, the relevant differences and similarities are a function of group membership and identity.⁷⁹ So I ought to be respected, not just because I am a human being, but also in virtue of my being a woman, or Jewish, or Zulu, or deaf etc. Not recognising and respecting my group identity is not recognising and respecting *me*. Supporting this view is a theory about the importance of group membership for identity formation.⁸⁰ Some judgments in *Barnard* recognise this shift to group membership; Van der Westhuizen J says that affirmative measures may affect the 'right to human dignity of people, individually or as members of a group'.⁸¹

'Equality' can be used for both individuals and for groups, as well as for individuals *qua* members of a group. That is, the following claims are possible (even if one does not agree with each): individuals are legally (and morally) equal to each other; socially and politically relevant groups are legally (and perhaps morally) equal; and individuals *qua* members of groups are legally (and morally) equal. Usually in *Barnard*, and in discussions over affirmative action more generally, 'equality' refers to persons *qua* members of a group (or sometimes to groups themselves). South Africa is trying to achieve equality for the group

⁷⁸ The classic essay on this vast topic is C Taylor 'The Politics of Recognition' in A Gutman (ed) *Multiculturalism: Examining the Politics of Recognition* (1994).

⁷⁹ The group memberships that are taken to be relevant vary. 'Identity Politics' is influential in feminism and race theory, where being a woman (belonging to the group 'women') or being black, is politically and morally relevant. See IM Young *Justice and the Politics of Difference* (1990). Other identity-conferring group memberships discussed in the ever-expanding literature are cultural, religious, political, ethnic etc.

⁸⁰ Taylor's essay gives an influential account of the identity-conferring role of group membership and why that membership should be respected: identity is not created *ex nihilo*, but 'dialogically'. 'We become full human agents, capable of understanding ourselves, and hence of defining our identity, through our acquisition of rich human languages of expression' (Taylor (note 78 above) at 32). These means of expression are acquired through our cultures. So we are human agents only in so far as we grow up in cultures that provide us with possible life narratives, and the expressive tools to articulate them. Our cultural membership is therefore a component of personal identity, and if persons are to be respected, then the conditions of their identity, among which will be cultural membership, should be supported. Will Kymlicka gives an influential liberal defence of this position (against so-called 'communitarian' positions) in W Kymlicka *Multicultural Citizenship: A Liberal Theory of Minority Rights* (1995).

⁸¹ *Barnard* (note 2 above) at para 168.

‘black people’, whose members were afforded unequal status during apartheid, and who still experience entrenched patterns of disadvantage, so a black person might receive particular treatment as a member of the group ‘black’. However, sometimes the individual claim is also put in terms of equality in *Barnard*, and then this is set up against the equality of a group, like the previously disadvantaged, or more vaguely, ‘society’. For example, Cameron and Froneman JJ and Majiedt AJ say that ‘there is a tension between the equality entitlement of an individual and the equality of society as a whole’.⁸²

I take it that they have legal and socio-economic equality in mind here, but we can recast equality talk into dignity talk, if equality is a necessary component of respecting dignity, or a necessary means towards protecting it. In this case, Ms Barnard’s ‘equality entitlement’ is her entitlement to equal consideration for jobs and promotions for which she is qualified. This talk of ‘individual’ entitlement to *equality* is, however, not as common in *Barnard* as talk of the entitlement of each individual person to have her *dignity* respected. That is, most often what is at stake is Ms Barnard’s dignity, not her equality, and at times equality of persons seems to collapse into the dignity of persons. So, ignoring the debate about which value, if any, is more fundamental in the Constitution, I shall for the most part be concerned with Ms Barnard’s dignity, and save ‘equality’ for the socio-political equality for the group of those disadvantaged by apartheid and its legacy.

In the context of the case, Ms Barnard’s dignity on the whole depends on her abilities and talents being properly recognised and taken into consideration. Even if correct, this view is not obvious, especially if dignity is given a Kantian interpretation, which makes it a function of a rational nature that we share. Just as a story needs to be told about the importance of group membership to identity, a story needs to be told about the role of abilities, talents and competencies for identity and for respectful treatment. It would presumably go something like this: if respecting the dignity of each person means respecting *her*, all that makes her the particular person she is, then we must properly acknowledge her abilities and professional merit. If she is the ablest at a job, she should be appointed to the job at the cost of not respecting her dignity. Note, however, that this account picks out features *particular* to Ms Barnard; it does not pick out her shared humanity, the basis for her dignity. The thought must be that one can only respect the shared humanity in a particular person by respecting the reason-infused capacities that humans have, and respecting how they are instantiated and exercised in different ways in different people.⁸³ Knowing how to respect Julia, and knowing how to respect James, will require knowledge and respect for the different rationally-infused qualities that make Julia and James morally distinct people, or the different ways in which they each realise and express them. However, a person is far more than her abilities, and there is a sense in which dignity is not fully respected if she is seen solely in that light. (I will return to this point.)

That is one aspect of respecting dignity. On the other hand, Ms Barnard’s membership of the advantaged group ‘white’ is just as important in this case

⁸² Ibid at para 77.

⁸³ A good account of how reason infuses many human activities is found in T Metz *Meaning in Life* (2013) ch 12.

as her status as a rational human being, and the particularity expressed in her talents and abilities. The judgments, especially the judgment of Cameron and Froneman JJ and Majiedt AJ, think it necessary to see Ms Barnard ‘through the lens’⁸⁴ of her group membership and that can both undermine and enhance her dignity-based claims. First, she is, relevantly, a woman. Gender is one of the ‘grounds’ on which a person cannot be unfairly discriminated against, but the gender ‘woman’ is also a ‘designated group’ which has been discriminated against in the past and so is deserving of redress now, as part of the goal of achieving equality. Second, however, she is white. White people continue to benefit from apartheid, and racial patterns of inequality still exist, so discriminatory measures are therefore required to make whites and blacks more equal (presumably by levelling up, not down). The judges, rightly, see that this intersection of group identities complicates the case. Cameron J, Froneman J, and Majiedt AJ note the failure of the National Commissioner to consider gender representivity in his reasons, and think that points to unfairness.⁸⁵ Van der Westhuizen J claims that because Ms Barnard’s ‘traits sit at the intersection of privileged and under-privileged identities, she might suffer harm in unique ways compared to members of other groups, designated or not’.⁸⁶ Her dignity might already be undermined by her being a member of the group ‘women’, and then ignoring her merit for the job is an added insult. And yet she is also a member of the group ‘white people’, and blacks have a claim for restitution against that group. Ignoring that claim is ignoring *their* dignity.

So far, there are four features of Ms Barnard’s identity that are relevant: her value (her dignity) simply as a human being; her particular value as the particular human being she is, with particular competencies; her group identity as a woman (and disadvantaged); and her group identity as white (and privileged). The first three of these are in her favour (as a human being; as *herself*; as a woman); the fourth (her being white) is the feature that gives rise to the tension. The fact that she is white means that other peoples’ dignity and the goal of social equality give the judges reason to weigh equality more heavily, but doing so risks disrespecting her dignity, and her disadvantaged position as a woman. Notice that *one* feature of Ms Barnard provides reason for weighing equality more heavily than the three other features, separately or combined. Also notice that the dignity of the majority of people (respected through equality measures) outweighs the dignity of one. That provides a straightforward consequentialist reason for setting aside her dignity in this case; numbers count, and numbers count especially once race is taken into account. Still, as we have seen, the judges do not rest content with such reasons; they try to be fair in a non-consequentialist sense in the way they come to their decision. That is, the numbers at stake are not deemed sufficient for justifying the decision, and for making the decision a fair one; as explored earlier, other criteria seem to be playing a role, and some of those are non-consequentialist.

⁸⁴ *Barnard* (note 2 above) at para 117.

⁸⁵ *Ibid* at para 120.

⁸⁶ *Ibid* at para 153.

One of those criteria, recall, was that the affected person be given reasons that she can understand and that make reference to principles and values she already accepts. Yet, there is a complication that comes out particularly in the case of white people in South Africa, who are officially the 'losers' in affirmative action (though of course that way of putting it is tendentious). In order for Ms Barnard (and any white person) to find the reasons behind the judgment acceptable, they must have accepted a number of things about themselves. One important fact whites have to acknowledge is that their whiteness *is* reasonably a mark against them. Defending this statement would take another paper: here I simply assert it.⁸⁷ In order for whites not to feel that they are being *unfairly* discriminated against in the very same way that blacks were under apartheid, they have to identify with, and fully acknowledge, their whiteness, and they have to acknowledge the negative meanings of whiteness. Whites have been especially adept at not considering their race important (when they acknowledge that they are raced at all); they often see it as incidental, not identity-defining. This nonchalance is impossible for blacks; in a white supremacist world, their race always matters. Whites therefore need to accept that in a world like ours their race is a legally and morally relevant fact about them, which they can only ignore in bad faith. Only once they have acknowledged this, will the judges' decision seem reasonable to them. Until then, they will feel unfairly discriminated against; they will feel that their dignity is being traded off against others' dignity in an objectionably crude consequentialist way, and, perhaps, that this is done intentionally as some kind of punishment for apartheid. These worries are certainly often raised in bad faith, but there is still something right in them: part of Kant's point is that one person cannot be sacrificed for many; that is constitutive of being 'an end in oneself'. The rights in the Constitution are our way of acknowledging this special value and are meant to rule out consequentialist reasoning at certain crucial points. I therefore do not think such worries are unjustified; there is something about affirmative action measures that remains ethically problematic even if in particular cases they are overall justified.

Earlier, I said that a person is far more than his abilities, and noted that perhaps dignity is not fully respected if he is seen only in their light. This thought is also lurking beneath this case. We all want to be valued 'for ourselves' and that vague phrase usually means more than 'the sum of our attributes'. There is a sense in which a person is instrumentalised if he is seen only as a bundle of capacities and talents and physical attributes, which could be put to use for service delivery and equality, or stand in their way. In affirmative action cases like these, a person is measured – and he *is* measured – by his usefulness, his group membership, his history, and – most apposite in South Africa – his race. That this is done in aid of restitution can feel hollow when the discrimination is committed in the very racialised terms used by the apartheid system. This assessment of each other is obnoxious; we want to be seen and valued as being more than our most visible or most politically relevant features; we want to relate to each other directly, not

⁸⁷ I have defended this in other work, eg Vice 'How Do I Live in 'This Strange Place?'' (note 6 above), and Vice 'Race, Luck, and the Moral Emotions' (note 6 above).

through the mediation of race. That this can sometimes be impossible, and that white people might have contributed to its being impossible, is cold comfort.

Finally, in order for (white) people to accept verdicts like the Court's in *Barnard*, they have to see socio-economic equality as a value in its own right, or as an indispensable condition for the dignity of those still disadvantaged. Just as they value their dignity, they should see that others value theirs as much⁸⁸ and that dignity requires socio-economic support; they should recognise that redress is called for even if they cannot welcome the sacrifice it requires of them – it is unreasonable to call on them to deny that it *would* be a sacrifice. This recognition could take a number of forms. First, they could buy into a shared South African political and social project, which would justify the sacrifice of some individual goods for the common good. Second, they could commit themselves to the ideal of moral equality, which would require that they commit themselves to improving conditions of those far less privileged than themselves, who are in their bad position because the group to which they belong was not treated equally. Third, they could learn to see their own wellbeing as dependent on the wellbeing of others – a notion drawn from the influential idea of Ubuntu. This idea is present in *Barnard*: indeed, Van der Westhuizen J shows his awareness of it when he alludes to a sense of community between rights holders, which could dilute the competition between claims. Each person, 'as the bearer of the right to dignity, should not be understood as an isolated and unencumbered being. Dignity contains individualistic as well as collective impulses',⁸⁹ he writes. This may be true, and the value of Ubuntu may indeed be informing the way the drafters of the Constitution understood the values of dignity and equality. However, exactly how this is to help in this case is not at all clear; perhaps it adds weight to the 'equality' side of the weighting; certainly, it asks us to see our dignity as dependent on the dignity of others. Without considerably more information, the addition of yet another value to the already complex set is a complicating, rather than a helpful factor.⁹⁰

The reference to Ubuntu does, however, alert us to the fact that in order to accept the reasons of the judges in *Barnard*, whites must think of themselves as part of one nation, a citizen amongst others with whom they are in mutual relations of dependency, trust and neighbourliness. They must already be committed to the goal of social equality that the Constitution sets out. Unless some kind of tie to others, *qua* South Africans, is felt and acknowledged, or unless they acknowledge basic moral solidarity with the victims of apartheid, they will feel they are the victims of unfair discrimination in cases like *Barnard*. They must already have made an ethical or political commitment to the goal of equality, and have accepted their implication in an unjust system, before affirmative action will seem just. Until this happens, they will feel as if their dignity is not being respected. Perhaps this ideal is what Van der Westhuizen J had in mind when he writes that

⁸⁸ This is a point Kant makes in a crucial passage in the *Groundwork*. Kant (note 23 above) at 4:429.

⁸⁹ *Barnard* (note 2 above) at para 174.

⁹⁰ Whether there is a difference and tension between dignity and Ubuntu in the South African Constitution is another matter of debate. See, eg, Cornell (note 57 above) and the response by Mokgoro & Woolman (note 57 above). There are different spellings of 'Ubuntu'/'uBuntu'.

‘restitutionary or affirmative measures should be welcomed rather than viewed with suspicion. They must be understood as equality-driven mechanisms in their own right, rather than carve-outs from what is discriminatory’.⁹¹

Perhaps this is a lot to accept, and bringing this case to the Labour Court in the first place is evidence that Ms Barnard did not fully accept it (though Van der Westhuizen J does commend her on her sensitivity).⁹² Now, to be clear, I *do* think she and other whites need to accept (most) of these, and I *do* think the judges’ verdict in *Barnard* is all-things-considered just and reasonable. So the difficulties I speak of here are not evidence that the judgment is not reasonable. The point is rather ethical and psychological: whites have to accept a lot of statements about themselves as true before the judges’ reasoning will resonate. Like all people, they have reasonable concerns for their own well-being and reasonable aspirations to be accepted on their own merits, and Van der Westhuizen J agrees that they will need to make a sacrifice:

[Ms Barnard’s] race was the determinative factor in the National Commissioner’s decision not to promote her. Her attributes, experience and attitude were eclipsed by considerations of race. Her value as a human being in an employment environment was, to some extent, undermined.⁹³

It is not helpful to expect any person to be unconcerned about her life, her security and her prospects and *happily* to give up the means for improving her condition. Affirmative action, even when its rationale is appreciated and accepted, is still something of a sacrifice for each person, and in part it undermines our moral system. This does not mean that it cannot be justified, only that its justification might be difficult to see if you do not already accept the goal.

IV CONCLUSION

Let me try to bring together the different strands of this paper. Insofar as there is an argument, one conclusion is that unjust circumstances can make it necessary to treat people in ways that can appropriately be seen as an affront to their dignity. As we see in *Barnard*, dignity will sometimes (perhaps often) need to be weighed against dignity and against other values. Trade-offs between social goals and individual dignity are inevitable. This does not, however, mean that they are made without moral costs. If Kant is correct and dignity has no price, then we are doing something wrong when we make these kinds of practically necessary calculations. Weighing one person’s dignity against others’ dignity or against equality is a *moral* mistake. You can do it, as we see in *Barnard*, but even if it is all-things-considered permissible, it can only be seen as the ‘least bad’ option, and this is already to accept the point that it *is*, indeed, bad. In an unjust world values have to be sacrificed and decisions made that are not clean, that have a moral residue. Equality and dignity are founding values in South Africa’s Constitution, but because of the facts of our socio-economic situation they will be impossible to realise together for a long time; and because of the complexity of

⁹¹ *Barnard* (note 2 above) at para 137.

⁹² *Ibid* at para 131.

⁹³ *Ibid* at para 177.

identity, many people will feel that their dignity is being impermissibly traded off. Comparing and weighing dignities *is* disrespectful, and in that regard, those like Ms Barnard would not be unreasonable if they felt undermined. I am not naïve enough to think we can do without trade-offs in this country, nor that examining the conflicts between values will stop us from comparing and weighing them. But there is some gain in integrity in facing up to the moral resonance of what we are forced by circumstances to do.

Despite the judges' commendable care and vigilance, then, *Barnard* suggests that the tension is irresolvable, at least in the current situation, and a harmonious reading of the Constitution's axiological commitments is therefore difficult to sustain.⁹⁴ We *cannot* have it all, both because of the facts of post-apartheid South Africa and because of the particular content and role given to the founding values in the Constitution. Perhaps this tension is contingent, and in better circumstances would not arise or would be resolvable without moral loss. However, in South Africa now, a clean solution is not possible, and this is a price we pay for apartheid.

However, and this is a second conclusion, there are different ways that these necessary but still bad trade-offs can be justified. Some will strike us as fairer than others and, as we saw, whether the decision against Ms Barnard was fair is a central issue in the judgments. I suggested that some of the criteria for fairness we can glean from the judgments are not very helpful, and I then suggested that the criterion of 'public reason' offers the best account of what makes a particular case of discrimination fair or unfair. This requires judges to make their reasoning transparent to the persons it affects, in terms that they can appreciate and where possible, share. In a Kantian framework, public reasons respect a person's rational nature, the source of her dignity. They give her reasons for a decision that negatively affects her in a manner that she can make sense of, that appeals to moral principles or values to which she sees the point or to which she is already committed. In that respect her dignity and moral status are appropriately acknowledged even if the particular decision is not to her liking. However, this position on fairness requires a lot more work to explain and defend it, which I cannot do here, and it is bound to be controversial in South Africa, with its uneasy mixture of liberal and traditional values.

A third conclusion, however, is that in order for whites to accept the reasoning of the judges as fair and public, they must already have undergone a significant (and required) moral change. They must have admitted that they continue to benefit from the ongoing unjust legacy of apartheid. They must have accepted the moral and political meanings of whiteness. They must have recognised that the dignity of others now requires sacrifices from them. Those sacrifices certainly matter morally and they have reason to feel unhappy. However, if they do not accept the judges' verdict in such cases, they are not being unjustly treated. The work they need to do on themselves before the reasons become, also, *their* reasons, *qua* citizens in a shared polity, therefore further complicates the publicity criterion for fairness. The judges' reasoning might indeed be suitably public, and so their verdict could be fair, without all those affected seeing this. This lack of insight on the part of those affected does not make the reasoning itself unacceptable;

⁹⁴ Ibid at paras 77–78 (Cameron J, Froneman J, and Majiedt AJ); at para 16 (Van der Westhuizen J).

they are in the wrong here, and their intransigence is part of the problems South Africa still grapple with. However, in a fractious country like ours, this situation makes reconciliation and racial justice more difficult.

Finally, there are two more general conclusions to draw from this discussion. The first is that there are real conflicts between different values and that a decision in favour of one does not mean a resolution of the conflict. We can see the conflict between dignity and equality in discrimination cases as a moral dilemma: any decision (and there must be a decision) will have some moral mark against it; there is no decision in which nothing will be sacrificed. The second general conclusion is that like so many difficult issues in politics and morality, we see a clash between two fundamental, and entirely different, approaches to morality. The Constitution is informed by the deontological approach of Kantian ethics, which requires that we respect people as ends in themselves at all times. That means we cannot view them solely as means to further social goals, and some treatment will be ruled out categorically. On the other hand, the needs and dignity of those people harmed by apartheid must be recognised and ameliorated. In this context, with scarce resources to share among millions of needy people, the sacrifice of an individual's dignity might be required, using practically necessary consequentialist calculations that might have results that are impermissible for a Kantian.

None of these conclusions offer solutions. If they are plausible, they are simply evidence of how difficult ethics and law is in a country in which ethics and legality were ignored for so long.

Affirmative Action and Intensity of Review: *South African Police Service v Solidarity obo Barnard*

Chris McConnachie*

I INTRODUCTION

In assessing the merits of an affirmative action measure, as with the merits of any law or conduct, a court must engage in at least two levels of reasoning.¹ At the first level, a court must consider the reasons for and against the measure; its substantive merits. At the second level, a court must determine the appropriate intensity of review. This involves deciding whether, and to what extent, it is institutionally appropriate for a court to interrogate the substantive merits. In practice, these first and second levels of reasoning are intertwined. However, it is important to recognise that they are conceptually distinct parts of a court's reasoning process, even if they are inseparable in application.

This article is concerned with this second level of reasoning: the process of determining the appropriate intensity of review in affirmative action cases. The Constitutional Court's decision in *South Africa Police Service v Solidarity obo Barnard* was widely expected to provide guidance and clarity on this issue.² The majority judgment did not do so, but the divergent views expressed by the judges are instructive. In this article, I use *Barnard* to explore how the Court's approach to the intensity of review ought to develop in future affirmative action decisions.

By the time the Court handed down judgment in *Barnard*, more than ten years had passed since *Minister of Finance v Van Heerden*.³ *Van Heerden* was the Court's first and only decision scrutinising an affirmative action measure for compliance with

* Advocate, Thulamela Chambers, Johannesburg Bar; Research Associate, Rhodes University. I would like to thank the participants at the Constitutional Court Review workshop in November 2015 for their valuable input. Special thanks to David Bilchitz and Khomotso Moshikaro for their editorial input and guidance and to the anonymous reviewers for their helpful comments and suggestions. I am grateful to Tarunabh Khaitan and colleagues at Oxford for the many conversations that helped to shape my arguments in this article.

¹ See further E Kavanagh 'Deference or Defiance? The Limits of the Judicial Role in Constitutional Adjudication' in G Huscroft (ed) *Expounding the Constitution: Essays in Constitutional Theory* (2009) 190 ('Deference or Defiance?'); E Kavanagh 'Defending Deference in Public Law and Constitutional Theory' (2010) 126 *Law Quarterly Review* 222, 230; C Chan 'Proportionality and Invariable Baseline Intensity of Review' (2013) 33 *Legal Studies* 1, 12 ('Baseline Intensity'). Both Kavanagh and Chan draw on J Raz *Practical Reason and Norms* (2nd Edition, 1999).

² *South African Police Service v Solidarity obo Barnard* [2014] ZACC 23, 2014 (6) SA 123 (CC), 2014 (10) BCLR 1195 (CC) ('*Barnard*').

³ *Minister of Finance & Another v Van Heerden* [2004] ZACC 3, 2004 (6) SA 121 (CC), 2004 (11) BCLR 1125 (CC) ('*Van Heerden*').

s 9(2) of the Constitution.⁴ That decision transformed s 9(2) from an interpretive guide into a fully fledged test for adjudicating the validity of affirmative action measures.⁵ The resulting ‘*Van Heerden* test’ was the Court’s first attempt to sketch the appropriate intensity of review in affirmative action cases.

In the decade after *Van Heerden*, lower courts largely ignored the Constitutional Court’s test.⁶ That much was evident as *Barnard* worked its way through the court system over a period of more than seven years. Three separate courts, including the Supreme Court of Appeal, failed to apply the *Van Heerden* test.

Given its path to the Constitutional Court, *Barnard* was tipped to be the Court’s moment to reassert and refine the *Van Heerden* test. The Court was also expected to clarify how the *Van Heerden* test should be tailored to the employment context and the specific provisions of the Employment Equity Act (EEA).⁷ In doing so, it was hoped that the Court would provide further guidance on the appropriate intensity of review.

The majority judgment and three separate concurrences in *Barnard* failed to live up to these expectations. Moseneke ACJ, writing for the majority dodged the central issues by holding that the unfair discrimination challenge was not properly before the Court. In passing, he affirmed that the *Van Heerden* test applies in assessing the validity of affirmative action measures under the EEA. However, he suggested, in obiter, that a different test is required in assessing the implementation of these measures, with rationality as its core. In his concurring judgment, Jafta J endorsed this rationality standard. The concurring judgment of Cameron, Froneman JJ and Majiedt AJ (Cameron et al) grappled with the unfair discrimination challenge. However, they did so by developing a new test for the validity of affirmative action under the EEA, one based on fairness. In a separate concurrence, Van der Westhuizen J was the only member of the court to apply the *Van Heerden* test in assessing the implementation of the affirmative action measure. The end result was a host of different approaches to the intensity of review with little guidance for future decisions.

In a detailed note on *Barnard*, Cathi Albertyn argues that the Court missed the opportunity to develop its affirmative action jurisprudence.⁸ In this article, I build on Albertyn’s analysis of the missed opportunities, focusing specifically on the intensity of review in applying the *Van Heerden* test.

⁴ Constitution of the Republic of South Africa, 1996. Section 9(2) states that: ‘Equality includes the full and equal enjoyment of all rights and freedoms. To promote the achievement of equality, legislative and other measures designed to protect or advance persons, or categories of persons, disadvantaged by unfair discrimination may be taken.’

⁵ The Constitutional Court is averse to using the term ‘affirmative action’, preferring the terms ‘remedial’ or ‘restitutionary’ measures (see *Van Heerden* (note 3 above) at para 29). However, I will refer to affirmative action as it is the term with the widest currency.

⁶ This has been a common pattern in the Labour Court and Labour Appeal Court. See further A Rycroft ‘Transformative Failure: The Adjudication of Affirmative Action Appointment Disputes’ in O Dupper & C Garbers (eds) *Equality in the Workplace: Reflections from South Africa and Beyond* (2009).

⁷ Act 55 of 1998.

⁸ C Albertyn ‘Adjudicating Affirmative Action Within a Normative Framework of Substantive Equality and the Employment Equity Act – An Opportunity Missed? *South African Police Service v Solidarity obo Barnard*’ (2015) 132 *South African Law Journal* 711.

I argue that the *Barnard* Court should have taken the opportunity to clarify the intensity of review in three respects. First, in support of Albertyn and other commentators, I argue that the Court should have affirmed that the *Van Heerden* test is applicable to affirmative action measures and their implementation under the EEA. Second, I also support the view that the Court should have clarified that the standard of proof embodied in the *Van Heerden* test is a proportionality analysis. Third, and most significantly, I go beyond the existing commentaries by exploring how and why the *Van Heerden* test may be applied with variable intensity.

Existing contributions to the South African literature focus on the standard of review to be applied in affirmative action cases.⁹ However, there has not yet been any serious engagement with the ways in which the chosen standard of review may be applied with different intensity depending on the context. As I will demonstrate, *Barnard* illustrates how this variable intensity of review will often be decisive. Given that so much turns on this variable intensity of review, it is an area of the Court's jurisprudence where far greater guidance and transparency is needed. In addition, this variable intensity of review is an important tool. It allows the Court to increase the intensity of review where this is necessary to prevent abuses and to come to the assistance of historically disadvantaged groups. I contend that the Court must openly justify its chosen intensity of review by reference to a set of three principles: the interests at stake, relative institutional competence, and considerations of democratic legitimacy. This is one of the central tasks for the Court to address in future cases.

I will develop this argument in five parts. Part II explains the nature and importance of the intensity of review in affirmative action cases. In Part III, I discuss the *Van Heerden* test and highlight where the test requires clarification and development. In Part IV, I provide some background to the Constitutional Court's decision in *Barnard*, setting out the facts and the lower courts' decisions. In Part V, I explore the Constitutional Court's reasoning in *Barnard* by assessing the majority and concurring judgments. I focus on the different standards of review adopted by the Court and the intensity with which these standards were applied. Finally, in Part VI, I set how the Court should have developed the *Van Heerden* test, emphasising the need for a principled approach to the variable intensity of review in future cases.

⁹ See, eg, C Albertyn & B Goldblatt 'Equality' in S Woolman & M Bishop (eds) *Constitutional Law of South Africa* (2nd Edition, 2008) 35-33-35-39; JL Pretorius 'R v Kapp: A Model for South African Affirmative Action Jurisprudence?' (2009) 126 *South African Law Journal* 398 ('R v Kapp'); JL Pretorius 'Fairness in Transformation: A Critique of the Constitutional Court's Affirmative Action Jurisprudence' (2010) 26 *South African Journal on Human Rights* 536 ('Fairness in Transformation'); JL Pretorius 'Accountability, Contextualisation and the Standard of Judicial Review of Affirmative Action: *Solidarity obo Barnard v South African Police Services*' (2013) 130 *South African Law Journal* 31 ('Standard of Review'); JL Pretorius 'The Unresolved Search for the Proper Standard of Review of Affirmative Action: *Solidarity obo Barnard v S.A.P.S*' (2013) 38 *Journal of Juridical Sciences* 128 ('Unresolved Search'); Albertyn *ibid*.

II INTENSITY OF REVIEW

A Importance and Substance

Before going further, it is necessary to say something about the importance of studying the intensity of review in affirmative action cases. On first glance, second-level reasoning about the appropriate intensity of review can appear formalistic and value-free. First-level reasoning about the substantive merits of affirmative action measures can often appear more urgent and important. First-level reasoning directly engages subjects such as the aims and values that should inform affirmative action measures, the effectiveness of these measures, and the balance between the benefits of specific measures and their costs.¹⁰ However, this does not mean that the intensity of review is any less significant or value-driven.

The United States Supreme Court's problematic affirmative action jurisprudence is a cure for any doubts about the importance and value-laden nature of the intensity of review. The Supreme Court's 'strict scrutiny' approach to race-based affirmative action has had profound effects, forcing most race-based affirmative action into hiding or retreat.¹¹ This standard of review sets a high bar by requiring any race-based measure to serve a compelling purpose and to be the only means of achieving that purpose. The ongoing debates on the Supreme Court over strict scrutiny also show that the choice of intensity of review is driven by divergent values.¹² The liberal and conservative judges' views are based on fundamentally different assumptions about racial injustice, the importance of affirmative action, and the role of the courts in evaluating these measures.

Fortunately, the South African Constitutional Court is not heading in the direction of a restrictive, US-style jurisprudence.¹³ The Court has rightly distanced itself from the strict scrutiny approach, holding it up as a benchmark of failure.¹⁴ Nevertheless, the challenge of finding an appropriate judicial role in evaluating affirmative action in South Africa is no less important.

While our courts are not chafing against the restraints of an overly restrictive affirmative action jurisprudence, they are faced with the opposite problem. They currently lack appropriate guidance on the appropriate intensity of review in these cases. *Barnard* and the string of other affirmative action cases currently before the

¹⁰ These issues have been canvassed extensively in the South African literature. See, eg, O Dupper 'In Defence of Affirmative Action in South Africa' (2004) 121 *South African Law Journal* 187; O Dupper 'Affirmative Action: Who, How and How Long?' (2008) 24 *South African Journal on Human Rights* 425; P de Vos 'The Past is Unpredictable: Race, Redress and Remembrance in the South African Constitution' (2012) 129 *South African Law Journal* 73.

¹¹ For an overview of this jurisprudence, see S Fredman *Discrimination Law* (2nd Edition, 2011) 118–125; S Fredman *Comparative Study of Anti-Discrimination and Equality Laws of the US, Canada, South Africa and India* (2012) 64–65; R Siegel 'Equality Divided' (2013) 127 *Harvard Law Review* 1.

¹² See, eg, the debates between the majority and minority in *Adarand Constructors Inc v Peña* 515 US 200 (1995); *Grutter v Bollinger* 539 US 306 (2003); *Fisher v University of Texas* 570 US – (2013); *Schnette v Coalition to Defend Affirmative Action, Integration and Immigration Rights and Fight for Equality By Any Means Necessary* 572 US – (2014).

¹³ For a comparison of the South African and US approaches, see K Naff & O Dupper 'Footprints Through the Courts: Comparing Judicial Responses to Affirmative Action Litigation in South Africa and the United States' (2009) 25 *International Journal of Comparative Labour Law and Industrial Relations* 157.

¹⁴ See *Van Heerden* (note 3 above) at paras 29, 147.

courts show that this uncertainty is fertile ground for litigation, particularly for groups seeking to frustrate affirmative action measures.¹⁵

There is no escaping the fact that the task of developing an appropriate intensity of review is a complex, technical exercise. That does not make it formalistic, in the sense of being unmoored from values and deeper debates about the merits of affirmative action.¹⁶ However, it remains a task laden with complexity. Part of the complexity is conceptual, as the intensity of review has many moving parts. In what follows, I briefly explain the intensity of review, its connection with deference and the ways in which these concepts manifest in judicial decisions.

B Intensity of Review, Deference and Their Manifestations

The intensity of review broadly refers to the strictness with which a court assesses the validity of laws and actions. This is generally related to the degree of deference that a court shows to the decision-maker in each case. Timothy Endicott helpfully defines deference as a court's willingness to 'leave the answer to some question, to some extent, to the initial decision-maker'.¹⁷ It is also common to see deference described as the process of giving greater 'weight' to the decision-maker's reasons for a law or action than those reasons would otherwise deserve.¹⁸ Both descriptions capture the idea that deference involves courts suspending judgment on the substantive merits of a law or action, at least to some degree.

A court's intensity of review and its degree of deference are generally inversely related. The more deferent the court the less intensely it will scrutinise laws and conduct, and vice versa. However, this relationship does not always hold. At times a court may go so far as to supply its own arguments and evidence in favour of a law or action, going beyond what is presented by the parties. In doing so, the court is no longer deferring to the decision-maker, in the sense of leaving certain matters unquestioned or giving greater weight to reasons provided. Instead, it is assuming a more active role in defending the law or action. This is not necessarily wrong, but it is important to watch for slippage where the language of deference is used to mask a more proactive or even partisan approach.

A court can adjust the intensity of review in two primary ways: by using different *standards of review* and by varying the intensity with it *applies* these standards.¹⁹

¹⁵ See Albertyn (note 8 above) 712–713; S Budlender, G Marcus & N Ferreira *Public Interest Litigation in South Africa: Strategies, Tactics and Lessons* (2014) 16–17 (on the attempts by the trade union Solidarity to resist affirmative action).

¹⁶ On the distinction between formalism and substantive reasoning, see PS Atiyah & RS Summers *Form and Substance in Anglo-American Law: A Comparative Study of Legal Reasoning, Legal Theory, and Legal Institutions* (1987) 1, 2, 5.

¹⁷ T Endicott *Administrative Law* (2nd Edition, 2011) 225.

¹⁸ Kavanagh 'Deference or Defiance?' (note 1 above) at 185; C Chan 'Deference, Expertise and Information-Gathering Powers' (2013) 33 *Legal Studies* 598, 600 ('Expertise').

¹⁹ For further discussion of this distinction see J Rivers 'Proportionality and Variable Intensity of Review' (2006) 65 *Cambridge Law Journal* 174 ('Variable Intensity'); Chan 'Baseline Intensity' (note 1 above) 5. Some have questioned the value of distinguishing between different standards of review, suggesting that it would be simpler to ask whether there are compelling reasons for a law or action in each context. That debate is beyond the scope of this article, but it suffices to say that the different standards of review have the benefit of providing analytical guidance, as opposed to a free-wheeling injunction to analyse the strength of the reasons.

Standards of review are the different sets of questions that a court asks in reviewing a law or action.²⁰ These standards of review exist on a spectrum, ranging from higher to lower levels of intensity. At the more intense end of this spectrum are standards of review such as the US Supreme Court's strict scrutiny approach, requiring a compelling purpose for race-based affirmative action measures and for these measures to be the only means of achieving this purpose.²¹ On the lower end of the spectrum are standards of review like rationality, which assesses whether the law or action is rationally connected to some legitimate purpose.²² Rationality is now the Canadian Supreme Court's favoured standard of review in assessing affirmative action measures, as set out in its 2008 decision in *R v Kapp*.²³ Proportionality and its more loosely defined cousin, reasonableness,²⁴ occupy a space in between strict scrutiny and rationality, involving a balancing of the benefits and harms of laws and actions.²⁵

Choosing a standard of review is not the end of the matter. Each of the standards of review can be applied with variable intensity from case to case. As Julian Rivers points out, merely talking about different standards of review fails to capture this variability.²⁶ For example, a court can subtly adjust the proportionality analysis in many ways, placing a thumb on the scales.²⁷ Similarly, the rationality analysis can also be applied more or less stringently, as commentators on the Court's rationality jurisprudence have repeatedly noted.²⁸

The intensity of review in applying these standards is intimately linked with the burden of proof and a court's willingness to play a proactive role in uncovering

²⁰ See generally S Seedorf & S Sibanda 'Separation of Powers' in S Woolman & M Bishop (eds) *Constitutional Law of South Africa* (2nd Edition, 2008) 12-59ff (OS 06-08).

²¹ See further Fredman (note 11 above) 118-125.

²² See, eg, *Prinsloo v Van der Linde and Another* [1997] ZACC 5, 1997 (3) SA 1012 (CC) at para 26 (differentiation must be rationally connected to a legitimate government purpose). Rationality has acquired other dimensions in the Court's s 1(c) rationality jurisprudence.

²³ *R v Kapp* [2008] 2 SCR 483 ('*Kapp*') (Canadian Supreme Court dismissed a challenge to a law allowing First Nations people to be given the exclusive licence to fish for salmon at valuable fishing grounds for a 24-hour period).

²⁴ On the relationship between proportionality and reasonableness, see, eg, C Steinberg 'Can Reasonableness Protect the Poor? A Review of South Africa's Socio-Economic Rights Jurisprudence' (2006) 123 *South African Law Journal* 264, 278-280; C Hoexter *Administrative Law in South Africa* (2nd Edition, 2013) 343-350 ('*Administrative Law*').

²⁵ Reflected in the Court's account of proportionality in *S v Manamela and Another (Director-General of Justice Intervening)* [2000] ZACC 5, 2000 (3) SA 1 (CC), 2000 (5) BCLR 491 (CC) at para 32 ('Court must engage in a balancing exercise and arrive at a global judgment on proportionality and not adhere mechanically to a sequential check-list').

²⁶ Rivers (note 19 above) 202ff.

²⁷ See, eg, the debate between the majority and minority in *Prince v President of the Law Society of the Cape of Good Hope and Others* [2002] ZACC 1, 2002 (2) SA 794 (CC), 2002 (3) BCLR 231 (CC) (Disagreement over the suitability and necessity of criminalising marijuana without exception for religious use).

²⁸ See C Hoexter 'Just Administrative Action' in I Currie & J De Waal (eds) *The Bill of Rights Handbook* (6th Edition, 2013) 688ff; M Bishop 'Rationality is Dead! Long Live Rationality! Saving Rational Basis Review' in S Woolman & D Bilchitz (eds) *Is This Seat Taken? Conversations at the Bar, the Bench & the Academy* (2012); A Price 'Rationality Review of Legislation and Executive Decisions: Poverty Alleviation Network and Albuti' (2010) 127 *South African Law Journal* 580; M du Plessis & S Scott 'The Variable Standard of Rationality Review: Suggestions for Improved Legality Jurisprudence' (2013) 130 *South African Law Journal* 597; L Kohn 'The Burgeoning Constitutional Requirement of Rationality and the Separation of Powers: Has Rationality Review Gone Too Far?' (2013) 130 *South African Law Journal* 810.

evidence and argument. Deference to the state is often shown in the way that courts give the state the benefit of the doubt. Judicial notice – the judicial recognition of facts that are considered to be general knowledge or are easily ascertainable from sources of incontrovertible authority²⁹ – also plays a role here. Courts may invoke judicial notice to fill gaps in the decision-maker's evidence. The court may also raise arguments or issues not addressed by the decision-maker as a way to bolster the reasons for or against a law or action.

In the existing South African literature on affirmative action, commentators have focused almost exclusively on the appropriate standards of review.³⁰ There has not yet been any sustained analysis of how and why these standards of review may be varied depending on the context. As I will argue in Parts V and VI, this is a crucial area for further development of the Court's affirmative action jurisprudence.

III THE *VAN HEERDEN* TEST

The Constitutional Court's decision in *Van Heerden* was its first opportunity to grapple with the appropriate intensity of review in affirmative action cases. Before *Van Heerden*, the Court approached affirmative action measures through the lens of the prohibition of unfair discrimination under s 9(3) of the Constitution.³¹ The validity of these measures was tested under the 'Harksen test' for unfair discrimination,³² developed in *Harksen v Lane* NO.³³

²⁹ PJ Schwikkard & SE van der Merwe 'Judicial Notice' in PJ Schwikkard & others (eds) *Principles of Evidence* (4th Edition, 2015) ch 27.

³⁰ See note 9 above.

³¹ Section 9 of the Constitution, the equality clause, provides that:

- (1) Everyone is equal before the law and has the right to equal protection and benefit of the law.
- (2) Equality includes the full and equal enjoyment of all rights and freedoms. To promote the achievement of equality, legislative and other measures designed to protect or advance persons, or categories of persons, disadvantaged by unfair discrimination may be taken.
- (3) The state may not unfairly discriminate directly or indirectly against anyone on one or more grounds, including race, gender, sex, pregnancy, marital status, ethnic or social origin, colour, sexual orientation, age, disability, religion, conscience, belief, culture, language and birth.
- (4) No person may unfairly discriminate directly or indirectly against anyone on one or more grounds in terms of subsection (3). National legislation must be enacted to prevent or prohibit unfair discrimination.
- (5) Discrimination on one or more of the grounds listed in subsection (3) is unfair unless it is established that the discrimination is fair.

The Court has not yet decided an unfair discrimination challenge based on s 9(4) of the Constitution.

³² See, eg, *President of the Republic of South Africa and Another v Hugo* [1997] ZACC 4, 1997 (4) SA 1 (CC), 1997 (6) BCLR 708 (CC) (Court found that the President's decision to pardon mothers of young children was not unfairly discriminatory); *City Council of Pretoria v Walker* [1998] ZACC 1, 1998 (2) SA 363 (CC), 1998 (3) BCLR 257 (CC) (Court found that a municipality's debt collection policy unfairly discriminated against white residents). It is arguable that in both cases, the measures would not have passed the *Van Heerden* test and would still have needed to be considered under the *Harksen* test. See further Albertyn & Goldblatt (note 9 above) 32–35.

³³ *Harksen v Lane NO and Others* [1997] ZACC 12, 1998 (1) SA 300 (CC), 1997 (11) BCLR 1489 (CC) at para 53 ('*Harksen*') (Court synthesised principles from earlier case law in setting out a structured test for unfair discrimination).

Section 9(2) of the Constitution merely served as an interpretive guide to this analysis.³⁴

Van Heerden changed this by stipulating that affirmative action measures should first be considered for compliance with s 9(2). If a measure satisfies s 9(2) then it is immunised from further scrutiny under s 9(3).³⁵ Only if it fails the *Van Heerden* test should it then be subjected to the *Harksen* test.

In creating the *Van Heerden* test, the Court was reacting to certain features of the *Harksen* test which required inappropriate forms of scrutiny for affirmative action. As a result, it is important to understand the *Harksen* test before considering the *Van Heerden* test in greater depth.

A From *Harksen* to *Van Heerden*

The *Harksen* test for unfair discrimination has two stages. First, a court must determine whether there is discrimination, which involves the imposition of burdens or the withholding of benefits on grounds listed in s 9(3) of the Constitution, or grounds analogous to these listed grounds. Second, if there is discrimination, then it must be determined whether the discrimination is unfair.

The unfairness analysis in the *Harksen* test is a complex and under-analysed part of the Court's jurisprudence. What is clear is that it involves two sets of enquiries: a) an analysis of the impact of the discrimination on the disfavoured group and b) an analysis of the justification for the discrimination, taking into account its impact.³⁶ This unfairness analysis is generally applied with varying intensity of review, as I have analysed in previous work.³⁷ The Court's standard of review in assessing the justification for discrimination often fluctuates between a rationality analysis and a proportionality analysis. There is also a great deal of variability in the intensity with which these standards of review are applied.³⁸

The *Harksen* test also includes a built in presumption of unfairness. Section 9(5) of the Constitution requires that, where discrimination has occurred on grounds listed in s 9(3), this discrimination must be presumed to be unfair, placing the

³⁴ See *National Coalition for Gay and Lesbian Equality and Another v Minister of Justice and Others* [1998] ZACC 15, 1999 (1) SA 6 (CC), 1998 (12) BCLR 1517 (CC) at paras 61–62 (explaining that s 9(2) requires 'substantive and remedial equality').

³⁵ *Van Heerden* (note 3 above).

³⁶ See further S Woolman & H Botha 'Limitations' in S Woolman & M Bishop (eds) *Constitutional Law of South Africa* (2nd Edition, 2008) 34–34–34–35 (OS 07-06); R Krüger 'Equality and Unfair Discrimination: Refining the *Harksen* Test' (2011) 128 *South African Law Journal* 479, 504–505; C McConnachie *What Is Unfair Discrimination? A Study of the South African Constitutional Court's Unfair Discrimination Jurisprudence* (DPhil Thesis, University of Oxford, 2014) ch 5.

³⁷ See C McConnachie 'Transformative Unfair Discrimination Jurisprudence: The Need for a Baseline Intensity of Review' (2015) 31 *South African Journal on Human Rights* 504, 516–524.

³⁸ *Ibid.*

burden squarely on the person alleged to be engaging in discrimination to prove otherwise.³⁹

The Court's efforts in *Van Heerden* to create a new test for assessing the validity of affirmative action should be seen against this backdrop.

B The Structure of the *Van Heerden* Test

Van Heerden concerned a challenge to a pension scheme for members of Parliament. The scheme paid bigger pension contributions to MPs who were elected after 1994. Those who served before 1994 were given a smaller contribution, reflecting the fact that the majority received generous pension packages under apartheid. The aim of this policy was to allow black MPs to develop pension savings comparable to their longer-serving, predominantly white colleagues. Frederick van Heerden, a white, National Party MP, challenged the scheme, arguing that it was unfairly discriminatory on the basis of race. The Court dismissed this challenge.

The key question facing the Court was how to go about evaluating this pension scheme. One option was to assess it under the s 9(3) prohibition of unfair discrimination, using the s 9(2) affirmative action provision merely as an interpretive aid. Another was to interpret s 9(2) as a standalone test for valid affirmative action measures, thus insulating valid affirmative action from further scrutiny under s 9(3) if it satisfies that test.⁴⁰

Moseneke J, writing for the majority, chose the latter option by interpreting s 9(2) as a standalone test. He held that affirmative action measures will only be scrutinised under the *Harksen* test if they fail the s 9(2) test.⁴¹ Moseneke J proceeded to outline three requirements for the validity of affirmative action measures, based on the wording of s 9(2).

First, the measure must '[target] persons or categories of persons who have been disadvantaged by unfair discrimination', requiring that members of historically disadvantaged groups should make up an 'overwhelming majority' of the beneficiaries of the measure.⁴²

Second, the measure must be 'designed to protect or advance such persons or categories of persons', meaning that it must be conducted for the purpose of benefitting disadvantage groups, it must not be 'arbitrary, capricious or

³⁹ In theory, this ought to require more stringent scrutiny in applying this analysis, although the Court has often wilfully ignored the s 9(5) presumption in practice. See, eg, *S v Jordan and Others* [2002] ZACC 22, 2002 (6) SA 642 (CC), 2002 (11) BCLR 1117 (CC) (Court found that the criminalisation of sex workers, but not their clients, is fair discrimination, despite the state offering no justification for the discrimination); *Volks NO v Robinson and Others* [2005] ZACC 2, 2005 (5) BCLR 446 (CC) (Court found that the exclusion of life partners from the Maintenance of Surviving Spouses Act 27 of 1990 was fair discrimination, despite the state and the executor conceding that the law was unfairly discriminatory).

⁴⁰ Pretorius 'R v Kapp' (note 9 above) terms these the 'integrative' and 'exemptive' approaches.

⁴¹ *Van Heerden* (note 3 above) at para 36.

⁴² Ibid at paras 38, 40. This accommodates cases of indirect affirmative action, where measures are not explicitly targeted at particular disadvantaged groups, but have the purpose and effect of benefitting disadvantaged groups nonetheless. On the nature and merits of indirect affirmative action, see T Khaitan *A Theory of Discrimination Law* (2015) ch 8.

display naked preference', and it must be 'reasonably likely' to protect or advance historically disadvantaged people.⁴³

Third, the measures must 'promote the achievement of equality'.⁴⁴ Moseneke J suggested that this involves some consideration of the impact of the affirmative action measure on those who are excluded:

Determining whether a measure will in the long run promote the achievement of equality requires an appreciation of the effect of the measure in the context of our broader society. It must be accepted that the achievement of this goal may often come at a price for those who were previously advantaged. ... However ... a measure should not constitute an abuse of power or impose such substantial and undue harm on those excluded from its benefits that our long-term constitutional goal would be threatened.⁴⁵

In his commentary on the case, JL Pretorius has argued that the *Van Heerden* test verges on a rationality analysis.⁴⁶ That would be true if the Court had left the test at the first two requirements. However, the inclusion of the third requirement of the test clearly contemplates some form of proportionality analysis, which involves weighing up the benefits of the affirmative action measure against its impact on those who are excluded.⁴⁷ A court could only determine whether the harm to the excluded is sufficiently 'substantial' and 'undue' by engaging in an exercise of proportionate balancing.

C The Justification for a Separate Test

The *Van Heerden* test is not entirely distinct from the *Harksen* test. Both tests involve some balancing of interests and consideration of the impact of the measure. However, there are two important differences between these tests. These differences are motivated by problems in applying the *Harksen* test to affirmative action in the first instance.

The first problem in applying the *Harksen* test to affirmative action is the expressive harm of placing concern for privileged groups ahead of the needs of historically disadvantaged groups.⁴⁸ Under the *Harksen* test, a Court is required to focus on the impact of the discrimination on the disfavoured individual or group.⁴⁹ Given that privileged groups are the most likely to be excluded from affirmative action measures, the *Harksen* test would place their interests front-and-centre in the analysis. This risks suggesting that the benefits of an affirmative action measure for historically disadvantaged groups are only of secondary concern. The third requirement of the *Van Heerden* test still leaves room for an assessment of

⁴³ *Van Heerden* (note 3 above) at para 41–43.

⁴⁴ *Ibid* at para 44–45.

⁴⁵ *Ibid* at para 44.

⁴⁶ See, eg, Pretorius 'Fairness in Transformation' (note 9 above) 537, 561–567.

⁴⁷ See J Brickhill 'Testing Affirmative Action under the Constitution and the Equality Act' (2006) *Industrial Law Journal* 2004, 2013; P de Vos (note 10 above) 93; Albertyn (note 8 above) 730–731.

⁴⁸ On the nature and significance of expressive concerns, see MD Adler 'Expressive Theories of Law: A Skeptical Overview' (2000) 148 *University of Pennsylvania Law Review* 1363, 1428–1447; T Khaitan 'Dignity as an Expressive Norm: Neither Vacuous Nor a Panacea' (2012) 32 *Oxford Journal of Legal Studies* 1, 5–9.

⁴⁹ *Van Heerden* (note 3 above) at para 80 (Mokgoro J noted this expressive problem).

the impact of affirmative action on disfavoured individuals and groups. However, it does not make this the focal point of the analysis, as the *Harksen* test does.

As a result, the *Harksen* and *Van Heerden* tests are different in the ordering of the impact analysis. The *Van Heerden* test first considers the benefits for the historically disadvantaged group and only then assesses its impact on those who are excluded or adversely affected. That ordering of the analysis may not have any practical significance, in the sense that this is unlikely to affect the outcome of decisions. However, it certainly has symbolic significance.

The second, and most significant, reason for this separate test is that the *Harksen* test would require courts to treat all affirmative action measures as presumptively unfair. That flows from the s 9(5) presumption of unfairness where discrimination occurs on grounds listed in s 9(3). Moseneke J held that it is inconsistent with the constitutional scheme to apply the s 9(5) presumption of unfairness in this way:

I cannot accept that our Constitution at once authorises measures aimed at redress of past inequality and disadvantage but also labels them as presumptively unfair. Such an approach, at the outset, tags s 9(2) measures as a suspect category that may be permissible only if shown not to discriminate unfairly. Secondly, such presumptive unfairness would unduly require the Judiciary to second-guess the Legislature and the Executive concerning the appropriate measures to overcome the effect of unfair discrimination.⁵⁰

This passage points both to expressive and institutional difficulties in holding affirmative action to be presumptively unfair.

The presumption is expressively problematic as it sends the message that all affirmative action is wrongful unless proved otherwise, rather than being a constitutionally sanctioned means of addressing patterns of group disadvantage. That is clearly inconsistent with the scheme of s 9 and the values that animate it.

The presumption of unfairness is also institutionally harmful. As Moseneke J acknowledges, this presumption may set the intensity of review too high. A presumption of unfairness may make courts all too willing to second-guess the state in complex matters, emboldening conservative litigants and judges to attempt to curtail affirmative action wherever possible. A further reason for rejecting this presumption is that it would require courts to err on the side of invalidating affirmative action measures where there is doubt about where the balance of probabilities lies. A presumption is, in essence, a technical device to allocate the risk of uncertainty.⁵¹ In the case of affirmative action, the potential for uncertainty is high. The effects of affirmative action, both positive and negative, will generally be apparent only after the measure has been in force over many years. In some cases, these effects may never be capable of accurate quantification. As Moseneke J acknowledged, the ‘future is hard to predict’.⁵² As a result, a presumption of unfairness would unduly curtail affirmative action, as it

⁵⁰ Ibid at para 33.

⁵¹ For more detailed analysis of presumptions in human rights litigation, see J Rivers ‘The Presumption of Proportionality’ (2014) 7 *The Modern Law Review* 409, 415–418 (‘Presumption of Proportionality’). See also C Chan, ‘The Burden of Proof under the Human Rights Act’ (2014) 19 *Judicial Review* 46 (‘Burden of Proof’).

⁵² *Van Heerden* (note 3 above) at para 41.

may often be impossible for the state to muster the necessary evidence at the time of litigation to prove that the benefits of particular affirmative action measures will outweigh the costs, on a balance of probabilities.

The Court undoubtedly had a further institutional worry in mind, although this is not openly articulated in *Van Heerden*. The presumption of unfairness raises the spectre that otherwise good and justifiable affirmative action programmes may be invalidated due to the state's failure to mount a proper defence in court. Even in cases where ample evidence and argument are available, the state has a track record of offering poor defences, or no defence at all. In such cases, the presumption of unfairness would not require the Court to 'second-guess' the other branches of state, as that presumes that there is reasoning before the Court to be scrutinised. Instead, the presumption would prevent the Court from coming to the defence of affirmative action measures by going beyond the meagre case presented by the state.

In his commentary on affirmative action decisions, Pretorius has suggested that the presumption of unfairness ought to be applied to affirmative action.⁵³ His argument appears to be that this presumption is a necessary component of a culture of accountability.⁵⁴ However, the need for accountability does not automatically translate into an argument for presumptive unfairness. Accountability requires that the state should, at a minimum, give reasons for its actions. But a duty of reason-giving is separate from the question of whether the state should be required to put up sufficient evidence and argument to overcome a presumption. A presumption sets a default position in the case of uncertainty, providing that affirmative action is unlawful unless proved otherwise. The symbolic and institutional difficulties of regarding all affirmative action as being unlawful by default are sufficient reason to reject the presumption in these cases.

The absence of a presumption of unfairness is the most significant difference between the *Van Heerden* and *Harksen* tests. The result is that the burden of proof will generally fall on the complainant to show that the affirmative action measure fails to comply with the s 9(2) requirements. However, the absence of a presumption of unfairness does not prevent a court from shifting the burden of proof to the opposing party in appropriate cases, either by tinkering with the burden of non-persuasion or the evidential and argumentative burden.⁵⁵ I will explore this important dimension of the variable intensity of review in greater detail in Part VI.

D Tasks for Future Decisions

The *Van Heerden* test left much work to do. Three tasks, in particular, were set for future decisions.

First, clarity was needed on whether the *Van Heerden* test applies to affirmative action measures under the statutory equality instruments, including the EEA and the Promotion of Equality and Prevention of Unfair Discrimination Act

⁵³ Pretorius 'Standard of Review' (note 9 above) 37–42.

⁵⁴ Ibid 39–40.

⁵⁵ I explain this distinction at text to notes 145–147 below.

(Equality Act).⁵⁶ Both statutes mirror s 9 of the Constitution, as they contain prohibitions on unfair discrimination, with the qualification that affirmative action is permissible.

Second, the precise standard of review embodied in the *Van Heerden* test also required development and clarification. All three requirements of the test were open-textured and required further refinement.

Third, the Court also needed to establish precisely how and why this standard of review may be applied with variable intensity. This variable intensity of review is an important and necessary component of the *Van Heerden* test. While affirmative action should not be treated as inherently suspect, there are likely to be cases where affirmative action measures may aggravate the patterns of disadvantage experienced by historically disadvantaged groups. For example, an affirmative action measure in the workplace that systematically excluded women would need to attract a heightened level of scrutiny. The variable intensity of review will allow the courts to respond appropriately in these situations. However, *Van Heerden* provided no guidance as to when and how this intensity of review may be adjusted.

It is no criticism of the *Van Heerden* Court to point to these loose ends. The Court made an admirable first step in developing a test while leaving room for development and invention. It was perhaps too much to expect the Court to offer a fully formed, precise test for such a contested issue in its first genuine affirmative action decision. The clear intention was that the test would gradually develop through use and exposure to real-world problems. However, a full decade would pass before the Court would have another opportunity to revisit the scope and content of the *Van Heerden* test.

IV BACKGROUND TO *BARNARD*

A The Facts

The facts in *Barnard* will now be familiar to most. Renate Barnard, a white woman, was a captain in the South African Police Service (SAPS). On two occasions, she applied for promotion to vacant positions. She was twice rejected despite being judged the best candidate. On both occasions, the interview panel recommended her for appointment. In her first application, her Divisional Commander declined to support her appointment as this would not advance racial representivity.⁵⁷ In the second application, it was the National Commissioner who declined her appointment, despite the Divisional Commander's recommendation.⁵⁸ The post was subsequently re-advertised and then scrapped in a process of restructuring.⁵⁹

The National Commissioner's reasons for rejecting Barnard's promotion were set out in a brief letter.⁶⁰ The National Commissioner reasoned that Barnard's appointment would not advance employment equity at the relevant salary level.

⁵⁶ Act 4 of 2000.

⁵⁷ *Barnard* (note 2 above) at para 9.

⁵⁸ *Ibid* at para 14.

⁵⁹ *Ibid*.

⁶⁰ *Ibid* at paras 15–16.

He also emphasised that the post was not critical and service delivery would not be affected if the post was left vacant until a suitable candidate could be found.

In reaching his decision, the National Commissioner was implementing the SAPS employment equity plan. The SAPS is a 'designated employer' for the purposes of the EEA.⁶¹ As a result, it is legally required to have an employment equity plan setting targets for the advancement of members of 'designated groups', defined under the Act as black people, women, and people with disabilities.⁶² The SAPS employment equity plan sets detailed numerical targets for the representation of members of designated groups at different salary levels. Barnard had applied for appointment at salary level 9, a level where white men and women were overrepresented.⁶³

The Employment Equity Act sets out requirements for how employment equity measures should be implemented. Section 15(1) provides that these measures should be 'designed to ensure that suitably qualified people from designated groups have equal employment opportunities and are equitably represented in all occupational levels in the workforce of a designated employer'. Section 15(3) further provides that while preferential treatment may be given to members of designated groups in pursuing the numerical goals set out in employment equity plans, these goals may not be treated as rigid quotas. Section 15(4) amplifies this by providing that affirmative action measures may not translate into absolute barriers to the employment or promotion of members from non-designated groups. The implication of these requirements is that the employment equity plan must be treated as a target rather than a set of rigid requirements.

In taking the decision to refuse Barnard's promotion, the National Commissioner was also implementing the 2004 National Instruction.⁶⁴ The instruction made the employment equity plan binding on all members of the SAPS by requiring interviewing panels to pay regard to the employment equity plan, among other considerations, in making decisions. All promotions to salary levels 8 and above had to be forwarded to the National Commissioner for his final determination. The instruction also specified that the fact that a candidate is judged to be the best does not require promotion. Furthermore, the National Commissioner could leave positions vacant if deemed appropriate.⁶⁵

After failing to resolve the matter in the CCMA, Barnard instituted proceedings in the Labour Court. She was represented by Solidarity, a conservative trade union which actively opposes affirmative action and has brought numerous cases in an attempt to advance its cause.⁶⁶ That fact could not have been lost on the courts.

In her statement of claim, Barnard alleged that she had been unfairly discriminated against on the basis of race under s 6(1) of the EEA.⁶⁷ Section 6(1)

⁶¹ EEA, s 1 defines designated employers as including 'an organ of state as defined in s 239 of the Constitution, but excluding the National Defence Force, the National Intelligence Agency and the South African Secret Service'.

⁶² EEA, s 1.

⁶³ *Barnard* (note 2 above) at para 66.

⁶⁴ *Ibid* at para 46.

⁶⁵ *Ibid*.

⁶⁶ See note 15 above.

⁶⁷ *Barnard* (note 2 above) at para 18.

largely mirrors the prohibition on unfair discrimination under s 9 of the Constitution.⁶⁸ Section 11 of the EEA also replicates s 9(5) of the Constitution by imposing a presumption of unfairness where discrimination occurs on listed grounds. However, s 6(2)(a) expressly provides for affirmative action, stating that 'it is not unfair discrimination to ... take affirmative action measures consistent with the purposes of the Act'. This provision had previously been interpreted as an interpretive guide in applying the test for unfair discrimination, rather than a standalone test for valid affirmative action.⁶⁹

In bringing her case to court, Barnard did not challenge the SAPS employment equity plan or the National Instruction. Her complaint was solely directed at the National Commissioner's implementation of these measures in deciding not to appoint her to the vacant post.

A further issue was the paucity of evidence either for or against the National Commissioner's decision. The National Commissioner did not testify in the Labour Court, nor did he depose to an affidavit explaining his reasoning.⁷⁰ It was left to a lower ranking official to testify in support of the National Commissioner's reasons, although he did not have personal knowledge of the decision.⁷¹ Barnard also made little attempt to counter the National Commissioner's reasons. In particular, she provided no concrete evidence to show that the National Commissioner's decision compromised service delivery.⁷²

As a result, the courts were presented with the difficult task of assessing the validity of the implementation of affirmative action measures, rather than the measures themselves, with little evidence on which to form an assessment. This combination of the narrowness of the challenge and limited information were to be decisive in the decisions that followed.

B Lower Courts' Decisions

Over a period of seven years, Barnard's case made its way from the Labour Court to the Labour Appeal Court and then to the Supreme Court of Appeal.⁷³ The courts reached different conclusions on the merits of her claim but their decisions shared a common failure to apply the *Van Heerden* test. They differed on

⁶⁸ Section 6 of the EEA provides, in relevant part:

- '(1) No person may unfairly discriminate, directly or indirectly, against an employee, in any employment policy or practice, on one or more grounds, including race, gender, sex, pregnancy, marital status, family responsibility, ethnic or social origin, colour, sexual orientation, age, disability, religion, HIV status, conscience, belief, political opinion, culture, language, birth or on any other arbitrary ground.
- (2) It is not unfair discrimination to
 - (a) take affirmative action measures consistent with the purpose of this Act; or
 - (b) distinguish, exclude or prefer any person on the basis of an inherent requirement of a job.'

⁶⁹ See note 6 above.

⁷⁰ *Barnard* (note 2 above) at para 104.

⁷¹ SAPS written submissions at para 74.

⁷² *Barnard* (note 2 above) at para 122.

⁷³ For commentary on these lower court decisions, see M Mushariwa 'Who Are the True Beneficiaries of Affirmative Action? *Solidarity obo Barnard v SAPS*' (2011) *Obiter* 439; Pretorius 'Standard of Review' (note 9 above); M McGregor 'Affirmative Action on Trial – Determining the Legitimacy and Fair Application of Remedial Measures' (2013) *Tydskrif vir die Suid-Afrikaanse Reg* 650; Pretorius 'Unresolved Search' (note 9 above).

what to make of the limited reasoning and evidence in support of the National Commissioner's decision.

In the Labour Court, Pretorius AJ applied the *Harksen* test in considering whether the National Commissioner's implementation of the employment equity plan breached s 6(1) of the EEA.⁷⁴ He found that the National Commissioner had discriminated against Barnard on the basis of race and that the discrimination had to be presumed to be unfair, in accordance with section 11 of the Employment Equity Act.⁷⁵ Given the absence of reasoning or evidence, Pretorius AJ concluded that the presumption had not been rebutted.⁷⁶

The Labour Appeal Court also applied the *Harksen* test, but concluded that the discrimination was fair.⁷⁷ In doing so, it reduced the unfairness analysis to a rationality enquiry.⁷⁸ In applying this standard of review, the LAC also appeared to place the burden of proof on Barnard. It concluded that the National Commissioner's decision was rationally linked to the employment equity plan.

The Supreme Court of Appeal reversed the LAC's finding.⁷⁹ In doing so, it also relied on the *Harksen* test and invoked the presumption of unfairness.⁸⁰ Having found that there was discrimination on the basis of race, the SCA held that the paucity of reasoning and evidence to support the National Commissioner's decision meant that the decision had to be presumed to be unfair.⁸¹

V BARNARD: THE CONSTITUTIONAL COURT'S DECISION

The confusion in the lower courts over the proper approach to affirmative action required a firm response from the Constitutional Court.⁸² It also gave the Court its first opportunity in a decade to reaffirm the *Van Heerden* test, to explain its application in the employment context, and to refine this test. In particular, clarity was needed on the precise standard of review required and intensity with which it ought to be applied.

The majority judgment and three concurrences were unanimous in the result: the National Commissioner's implementation of the employment equity plan was valid. However, the result was always going to be the least interesting and important part of the judgment. It was the manner in which the Court reached the result that was significant, as it was hoped this would provide guidance for future affirmative action cases. In this respect, the members of the Court were united by the view that it was impermissible to apply the *Harksen* test for unfair discrimination to the National Commissioner's decision. Beyond this point of

⁷⁴ *Solidarity obo Barnard and Another v South African Police Services* [2010] ZALC 10, 2010 (10) BCLR 1094 (LC).

⁷⁵ *Ibid* at para 26.

⁷⁶ *Ibid* at paras 36, 43.6.

⁷⁷ *South African Police Services v Solidarity obo Barnard* [2012] ZALAC 31, 2013 (3) BCLR 320 (LAC) at para 21.

⁷⁸ *Ibid* at para 44.

⁷⁹ *Solidarity obo Barnard v South African Police Service* [2013] ZASCA 177, 2014 (2) SA 1 (SCA).

⁸⁰ *Ibid* at paras 50 and 76.

⁸¹ *Ibid* at para 76.

⁸² *Barnard* (note 2 above).

agreement, the different judgments approached the issues in very different ways, applying different standards of review.

A The Mosenke Majority and Jafta Concurrence

Mosenke ACJ was the architect of the Court's *Van Heerden* test. It was expected that he would take the opportunity to apply this test under the EEA and, in doing so, that he would clarify and expand on the appropriate intensity of review. However, he largely avoided these issues in his majority judgment.

This avoidance involved three steps. First, Mosenke ACJ held that the SCA erred by applying the *Harksen* test for unfair discrimination at the outset, without first assessing whether the National Commissioner's decision amounted to the valid implementation of an affirmative action measure.⁸³ In doing so, the SCA had misapplied the law.

Second, Mosenke ACJ then declined to set out and apply the correct test for assessing the validity of the implementation of affirmative action measures. This was because he characterised Barnard's case as having abandoned the unfair discrimination challenge.⁸⁴ Instead, he held that her case had morphed into a quasi-administrative law review of the rationality and reasonableness of the National Commissioner's decision.⁸⁵ He suggested that this challenge must fail as it was not properly pleaded.⁸⁶

Third, he indicated that even if the Court were to consider the merits of the new challenge it would fail as the National Commissioner's decision had not been shown to be irrational and unreasonable.⁸⁷ This conclusion followed principally because Barnard bore the burden of proof.

Albertyn describes the majority judgment as an 'oddly formalistic side-step' that avoided the main issues.⁸⁸ The majority was correct that the basis of Barnard's challenge in the Constitutional Court was obscure. However, as Cameron et al made clear, the essence of the complaint had always been one of unfair discrimination.⁸⁹ This necessarily required the Court to engage with the validity of the implementation of the employment equity plan. Indeed, Mosenke ACJ acknowledged the essence of the case in the opening paragraph of his judgment, where he noted '[t]he core issue in [the SCA] and this Court remains unchanged. Did the National Commissioner's decision unfairly discriminate against the respondent?'⁹⁰

Seen in this light, the overwhelming impression created by the majority judgment is of a Court searching for an easy way out of a difficult task. In doing so, the majority did not merely miss the opportunity to develop and clarify the *Van Heerden* test, it actively declined this opportunity where it could and should have seized it. Of course, judicial minimalism has its place in constitutional

⁸³ Ibid at paras 51–53.

⁸⁴ Ibid at paras 58–59.

⁸⁵ Ibid.

⁸⁶ Ibid at para 60.

⁸⁷ Ibid at para 61–70.

⁸⁸ Albertyn (note 8 above) 716.

⁸⁹ *Barnard* (note 2 above) at para 82.

⁹⁰ Ibid at para 1.

jurisprudence, but not in this case.⁹¹ More was required of the Court given the confusion about the appropriate intensity of review in affirmative action cases.

While the majority judgment lays down very little precedent, it does establish one principle and suggests another. First, Moseneke ACJ confirmed that the *Van Heerden* test does apply under the EEA, but only in respect of assessing the merits of employment equity plans and other affirmative action ‘measures’, not their implementation.⁹²

Second, he suggested that the implementation of affirmative action measures is to be assessed by a different test.⁹³ He hinted that this test requires legality and rationality, at minimum, but he refused to be drawn on the precise standard of review or the intensity with which it is to be applied. He concluded ‘these are the minimum requirements, it is not necessary to define the standard finally’.⁹⁴ These comments on assessing the implementation of affirmative action were strictly obiter, but they may be developed in future decisions.

In his separate concurrence, Jafta J also strongly suggested that rationality is the preferred standard for assessing the implementation of affirmative action measures.⁹⁵ However, he argued that the question of the appropriate standard of review had not been raised by any of the parties.⁹⁶ This was not strictly correct, as the heads of argument prepared by the SAPS argued that the *Van Heerden* test should be applied in assessing affirmative action measures *and* their implementation.⁹⁷ This again suggests a retreat from deciding matters that were squarely before the Court.

B The Cameron, Froneman and Majiedt Concurrence

The judgment of Cameron et al criticised the majority for avoiding the key issues. In their view, the issue of unfair discrimination had been squarely raised and it fell to the Court to decide on the appropriate standard of review:

This is the first case before this Court that deals with the standard to be applied in assessing the lawfulness of the individual implementation of constitutionally compliant restitutionary measures. It is important to give guidance on this difficult issue.⁹⁸

⁹¹ On the appropriate place of minimalism in judicial reasoning, see CR Sunstein *One Case at a Time: Judicial Minimalism on the Supreme Court* (1999) ch 1. For further discussion of maximalism and minimalism in the South African context (and the related, but not identical, distinction between substantive and formalist reasoning), see, eg, A Cockrell ‘Rainbow Jurisprudence’ (1996) 12 *South African Journal on Human Rights* 1; I Currie ‘Judicious Avoidance’ (1999) 15 *South African Journal on Human Rights* 138; C Roederer ‘Judicious Engagement: Theory, Attitude and Community’ (1999) 15 *South African Law Journal* 486; S Woolman ‘The Amazing, Vanishing Bill of Rights’ (2007) 124 *South African Law Journal* 762; K Young ‘The Avoidance of Substance in Constitutional Rights’ (2015) 5 *Constitutional Court Review* 233.

⁹² *Barnard* (note 2 above) at para 38.

⁹³ *Ibid* at para 38–39.

⁹⁴ *Ibid* at 39.

⁹⁵ *Ibid* at paras 227–229.

⁹⁶ *Ibid* at para 219.

⁹⁷ SAPS written submissions at paras 63–68, see para 67 in particular.

⁹⁸ *Barnard* (note 2 above) at para 75.

Cameron et al proceeded to argue that the appropriate standard of review for assessing the implementation of affirmative action measures under the EEA is ‘fairness’.⁹⁹ In doing so, they appeared to agree with the majority that the *Van Heerden* test should not be applied to implementation. Precisely why they took this stance is not fully explained in the judgment. In a single paragraph, they suggest that unfair discrimination under the Constitution involves different considerations and problems to unfair discrimination under the EEA, necessitating a different approach to affirmative action.¹⁰⁰ However, that justification overlooks the fact that the *Harksen* test for unfair discrimination has been consistently applied under s 6(1) of the EEA, even though it was first developed under s 9(3) of the Constitution. There is no apparent reason why the *Van Heerden* test under s 9(2) cannot follow suit and be applied under the EEA, with some necessary tailoring to the context.

Cameron et al also found themselves with the difficult task of explaining how a test of ‘fairness’ in the context of affirmative action measures is different to a test for the fairness of discrimination. The same word can, of course, be used to describe different concepts.¹⁰¹ However, few will be so astute as to notice the difference, creating great potential for confusion.

The precise nature of the ‘fairness’ standard of review was not fully explained in the judgment. However, two features of this test were made clear. First, Cameron et al firmly indicated that ‘fairness’ involves more than rationality. In their view, rationality is too deferential to allow the courts to interrogate the balance between the benefits and harms of affirmative action measures. See, for instance, the following statement:

The important constitutional values that can be in tension when a decision-maker implements remedial measures require a court to examine this implementation with a more exacting level of scrutiny.¹⁰²

And further:

[A] rationality standard does not allow a court to interrogate properly a decision-maker’s balancing of the multiple designated groups, or of their interests against those adversely affected by the restitutionary measures.¹⁰³

We must therefore formulate a standard specific to the Act, one that is rigorous enough to ensure that the implementation of a remedial measure is “consistent with the purpose of [the] Act” – namely, to avoid over-rigid implementation, to balance the interests of the various designated groups, and to respect the dignity of rejected applicants.¹⁰⁴

Second, Cameron et al indicated that the more determinate rules on affirmative action in the EEA are bundled up in the fairness standard.¹⁰⁵ On this approach,

⁹⁹ Ibid at para 101.

¹⁰⁰ Ibid at para 97.

¹⁰¹ See, eg, *S v Dlamini* [1999] ZACC 8, 1999 (4) SA 623 (CC), 1999 (7) BCLR 771 (CC) at para 47 on the different meanings of ‘interests of justice’.

¹⁰² *Barnard* (note 2 above) at para 95.

¹⁰³ Ibid at para 96.

¹⁰⁴ Ibid at para 97.

¹⁰⁵ Ibid at para 100.

the prohibitions on quotas and job reservations under ss 15(3) and (4) of the EEA appear to fall within this all-things-considered fairness analysis.

In applying the fairness standard to the facts, Cameron et al gave some hint that this standard could be applied with variable intensity in different cases. The central question they confronted was whether the paucity of evidence and reasoning in support of the National Commissioner's decision was fatal. Cameron et al used a number of techniques to find that the balance ultimately weighed in the National Commissioner's favour.

First, the burden of proof was placed squarely on Barnard. It was clear from the way that Cameron et al considered the matter that the applicant bore the risk of non-persuasion and that the benefit of any doubt would be given to the National Commissioner.¹⁰⁶

Second, Barnard also carried the evidential and argumentative burden. Cameron et al emphasised that, in the circumstances, Barnard had to make out a case that her promotion was in the interests of service delivery, that gender representation would be advanced by her promotion, and that these interests outweighed the interests of promoting racial representation at salary level 9. The Court was not going to come to Barnard's assistance in making these arguments and providing evidence in support of them. However, Cameron et al did appear to suggest that if Barnard had based her case on gender rather than race, then the outcome could have been different.¹⁰⁷ Nonetheless, they held that '[i]n the absence of proper challenge and argument, the Court cannot undercut the decision-maker's stated reasons on this point'.¹⁰⁸

Finally, Cameron et al came to the assistance of the National Commissioner by going beyond the limited reasons he provided for the decision. While the National Commissioner failed to provide any real justification, Cameron et al held that 'this is not fatal, because there are sufficient external facts to determine that the National Commissioner's decision was fair'.¹⁰⁹ In this way, fairness can be determined on the objective facts, rather than the justification offered by the decision-maker. This approach is summed up in the final paragraphs of the judgment:

We conclude that the facts show that the National Commissioner's decision passes the fairness standard. While we find this is a close call, what has proved determinative to us is the pronounced over-representation of white women at the salary-level to which Ms Barnard was applying. This was not just one or two, but many. There was thus greater justification for prioritising racial representivity over other considerations.¹¹⁰

This approach is not only deferential to the National Commissioner, in the sense of affording greater weight to his reasons than they would otherwise carry. It involves the Court actively seeking out evidence and reasons in favour of the decision. This approach may well be justified in the circumstances of the case.

¹⁰⁶ Ibid at para 122.

¹⁰⁷ Ibid at para 115.

¹⁰⁸ Ibid at para 122.

¹⁰⁹ Ibid at para 104.

¹¹⁰ Ibid at para 123.

However, no explicit justification was given for applying the standard of review in this way.

The great point of uncertainty in this judgment is when and how the intensity of review may vary and for what reasons. Cameron et al were largely silent on this issue. At various points they signal that the way they have applied the evidential and argumentative burden may be different in other cases. In particular, they implied that if Barnard had made gender representation the focus of her case, then the Court would have applied a more intense level of scrutiny. They also strongly emphasised the importance of state accountability and reason-giving, suggesting that in appropriate circumstances the state's silence in defending specific instances of affirmative action may count against it.¹¹¹ However, no concrete guidance was offered for future decisions.

C The Van der Westhuizen Concurrence

In his separate concurrence, Van der Westhuizen J was the only member of the Court to apply the *Van Heerden* test in assessing the validity of the National Commissioner's decision.¹¹² In doing so, he also rejected the view that rationality suffices.¹¹³

While Van der Westhuizen J applied the *Van Heerden* test, he did not interpret it as incorporating a proportionality analysis. Instead, he suggested that this proportionality analysis is located outside the *Van Heerden* test.¹¹⁴ What is never explained in the judgment is why the balancing of interests cannot be done within the third leg of the *Van Heerden* enquiry, as the *Van Heerden* Court appeared to require.

In applying this analysis, Van der Westhuizen J proceeded to determine whether the National Commissioner's decision impacted on Barnard's dignity and undermined service delivery in a manner that was disproportionate.

Van der Westhuizen J placed the burden of proof squarely on Barnard. He also followed Cameron et al in considering evidence and arguments in favour of the National Commissioner's decision which had not been advanced by the National Commissioner himself. This approach is captured in his discussion of the difficulties of assessing the impact of the decision on service delivery:

Without proper evidence or specialist institutional knowledge, it may be difficult for a court to draw conclusions about the precise impact a policy, an appointment, or even a vacancy will have on service delivery. This is the reason for the National Commissioner's wide discretionary powers, particularly in the context of affirmative measures, to appoint a candidate or keep a post vacant. In this case there is not enough evidence for this Court to impugn the decision on the issue of service delivery. It cannot be said that it was disproportionate for the National Commissioner to rank representivity higher than the possible impact on service delivery in this case.¹¹⁵

¹¹¹ Ibid at paras 103–107.

¹¹² Ibid at paras 133, 142–143.

¹¹³ Ibid at para 141.

¹¹⁴ Ibid at para 165–167.

¹¹⁵ Ibid at para 189.

Like Cameron et al, there is little in the way of open justification for applying the proportionality analysis in this way. The justification is hinted at rather than addressed directly. That justification lies necessarily in the nature of the interests at stake, the court's institutional competence on these issues, and questions of the Court's legitimacy in this area. However these reasons were largely obscured by talk of dignity, equality and other values.

D Missed Opportunities

In sum, the *Barnard* Court did not provide definitive answers to the three questions raised by *Van Heerden*. This is likely to cause difficulties for courts confronted with affirmative action cases in future.

First, there is likely to be further confusion about the application of the *Van Heerden* test under s 6 of the EEA to the implementation of affirmative action measures. The one point on which the *Barnard* majority established precedent is that the *Van Heerden* test applies when considering the validity of affirmative action measures, such as employment equity plans. However, it remains uncertain whether the *Van Heerden* test applies to the implementation of these affirmative action measures. The majority judgment and the Cameron et al and Jafta concurrences all suggest that the implementation of affirmative action will be assessed by some other test, although no clarity is provided on what this should be.

Second, the divergent judgments offer no concrete guidance on the appropriate standard of review when assessing affirmative action measures. The majority judgment suggests an inclination to rationality, although this was not finally determined. Only Jafta J was firmly in favour of making rationality the exclusive standard of review. Cameron et al and Van der Westhuizen J favoured a proportionality analysis, involving the balancing of interests, although both located this analysis outside of the *Van Heerden* framework.

Third, there was no clear guidance on when and how these standards of review may be applied with varying intensity. This is significant as the judges' chosen intensity of review in applying their different standards proved to be the decisive factor in the case. Albertyn notes that it is surprising that the judges could reach the same result despite using such different standards of review.¹¹⁶ The result is not at all surprising, however, when we recognise that the majority and concurring judgments applied different standards of review with the same, low intensity. In particular, all of the judgments placed the burden of proof squarely on the applicant. Once the Court accepted that *Barnard* had the burden to prove that the implementation of the affirmative action measures was invalid, the result was a foregone conclusion, no matter what standard of review was applied. *Barnard*'s failure to put up real evidence or argument to challenge the National Commissioner's reasons, as scant as they were, meant that her claim could not succeed. The choice to place the burden of proof on *Barnard* was justified, for reasons I will expand on below. However, the members of the Court did not attempt to justify this choice or even acknowledge that there was a choice to be

¹¹⁶ Albertyn (note 8 above) 722.

made. That choice was central to the outcome of the case and required greater candour and explanation.

After a wait of ten years, *Barnard* did not deliver the expected guidance on these issues. The Constitutional Court did not, however, have to wait long for another affirmative action case. In its 2016 judgment in *Solidarity v Department of Correctional Services* the Court was again confronted with an affirmative action case; however the narrow scope of the arguments did not allow the Court to address the questions left open in *Barnard*.¹¹⁷

Solidarity concerned a direct challenge to the Department of Correctional Services' employment equity plan in the Western Cape. This was in contrast with *Barnard*, where the plan was not challenged. Ten prison warders, supported by *Solidarity*, challenged the Department's refusal to appoint them to various positions. Nine of the ten identified as Coloured and several were women. They were denied appointment on the basis that Coloured people and women were overrepresented in the relevant posts, according to the targets set in the employment equity plan. It was common cause that these targets were based on national demographics and did not take account of the specific demographics of the Western Cape. The Court found that s 42(a) of the EEA, as it was then worded, required due consideration of regional demographics.¹¹⁸ As a result, the Court held that the plan was in breach of s 42(a) and the refusal to appoint the warders on the basis of this flawed plan was therefore invalid.¹¹⁹

The Court was not required to reach any firm findings on the intensity of review in affirmative action cases. Nevertheless, the Court did lay down three important principles. First, the Court confirmed what it called the '*Barnard* principle': that the *Van Heerden* test applies when assessing the validity of affirmative action measures, including employment equity plans.¹²⁰ Second, the Court held that the *Van Heerden* test applies even when considering claims brought by historically disadvantaged individuals who have been denied benefits under an affirmative action scheme.¹²¹ This confirms that the *Van Heerden* test applies irrespective of the identity of the party challenging the affirmative action measure. Thirdly, the majority also rejected the argument that the employment equity plan was a disguised quota, which is prohibited under s 15(3) of the EEA.

Given the narrow scope of this decision, the Court did not engage further with the nature of the *Van Heerden* test and the appropriate intensity of review in applying this test. The Court also shed no further light on the appropriate standard of review in assessing the implementation of affirmative action measures, as it was primarily concerned with the employment equity plan itself

¹¹⁷ *Solidarity and Others v Department of Correctional Services and Others* [2016] ZACC 18, 2016 (5) SA 594 (CC), 2016 (10) BCLR 1349 (CC) ('*Solidarity*').

¹¹⁸ *Ibid* at para 74. At the relevant time, s 42(a) of the EEA provided that in assessing compliance with the EEA, the relevant bodies 'must' take into account the 'demographic profile of the national and regional economically active population'. Section 42(a) was subsequently amended to replace 'must' with 'may'.

¹¹⁹ *Ibid* at para 82.

¹²⁰ *Ibid* at para 25, summarising and endorsing the Labour Appeal Court's interpretation of *Barnard* in *Solidarity and Others v Department of Correctional Services and Others* [2015] ZALAC 6, 2015 (4) SA 277 (LAC) at para 51.

¹²¹ *Ibid* at paras 37–49.

and not its implementation. As a result, the questions left open in *Barnard* remain unanswered.

VI DEVELOPING THE *VAN HEERDEN* TEST IN FUTURE DECISIONS

I now turn to give arguments showing how the Constitutional Court can refine the *Van Heerden* test in future decisions. Three key developments are required, which I elaborate upon below.

A The Application of *Van Heerden* under the EEA

The first step, I contend, is that the Constitutional Court should apply the *Van Heerden* test both to affirmative action measures *and* their implementation.

It must be acknowledged that there is a conceptual and analytical distinction worth making between affirmative action measures and their application in practice. An employment equity plan may be valid on its face but may be applied in ways that are invalid. An abstract assessment of an employment equity measure may also involve different factual considerations than a detailed assessment of how the affirmative action measure is being applied in practice.

However, the fact that there is a difference between an affirmative measure and its implementation does not mean that different tests are required to assess their validity. There is no principled or practical reason for creating entirely separate tests.

First, as I noted earlier, it is difficult to see why two separate tests are needed for the validity of affirmative action measures and their implementation when one test suffices for unfair discrimination cases. There is no bifurcation of tests in assessing unfair discrimination challenges under section 6 of the EEA. The *Harksen* test is used to assess workplace policies and their day-to-day implementation. Precisely why the Court considered it necessary to have two separate tests for the validity of affirmative action measures was never adequately explained in *Barnard*.

Second, as Van der Westhuizen J pointed out, the *Van Heerden* test was clearly designed to consider both affirmative action measures and their implementation.¹²² This is evident in the third requirement, which focuses on the benefits of the affirmative action measure and its negative impact on those who are excluded. It would be impossible to consider the beneficial and harmful impact of affirmative action measures without any regard for how they are or will be implemented in practice. Assessing the validity of measures and their implementation in isolation would result in an impossible abstraction for a test which is meant to be contextualised in reality.¹²³

Third, and most importantly, separate tests may dilute the protections afforded to historically disadvantaged groups. This is because employment equity plans and their implementation both have the potential to harm disadvantaged groups to the same extent, and so require the same standard of review. For example, consider an employment equity plan that sets ambitious targets for the promotion

¹²² *Barnard* (note 2 above) at para 143.

¹²³ See *Van Heerden* (note 3 above) at paras 44, 142.

of disabled people but is implemented in ways that ignore the interests of this group. Compare this with an employment equity plan that, in its design, gives insufficient weight to the interests of disabled people. Both scenarios would likely impact and exclude disabled people in the same way. As a result, it would make no sense to subject the exclusionary plan to a more intense standard of review than the improper implementation of the otherwise inclusive plan. Doing so would risk allowing abuses and undue burdens to slip through without sufficient checks.

The application of the *Van Heerden* test under the EEA is not without challenges, which will require further thought as the case law develops. In particular, a clearer account is needed of how the *Van Heerden* test relates to the more determinate provisions in the Act. For example, the Act contains rule-like norms on regional representivity,¹²⁴ a prohibition on quotas and absolute job reservations,¹²⁵ and requirements that candidates be suitably qualified.¹²⁶ The judgments of Van der Westhuizen J and Cameron et al appear to suggest that these rule-like norms should be bundled into the broader balancing enquiry. This does not appear to be appropriate. It is the nature of rules that they are meant to be determinative of specific factual situations, providing more precise guidance to decision-makers.¹²⁷ Placing these rule-like norms into an all-things-considered balancing enquiry may dilute their force. As a result, it would be better to consider these rules as binding constraints that are independent of the *Van Heerden* test. Compliance with these rules is a necessary condition for the validity of an affirmative action measure, but it is not sufficient to guarantee such validity. An affirmative action measure that complies with these rules could still be invalidated under the *Van Heerden* test if it does not strike an appropriate balance between the competing interests at stake. For example, an affirmative action measure may not be a prohibited quota, but it may still fail the *Van Heerden* test.

B The Standard of Review

The next development that is needed is confirmation that the *Van Heerden* test involves a proportionality analysis, rather than a mere rationality test. As Albertyn notes, Van der Westhuizen J was on the right track in this respect.¹²⁸ However, he curiously decided to locate this proportionality analysis outside of the *Van Heerden* test. That is unnecessary as the third requirement of this test is sufficiently flexible to accommodate this analysis.

The judgment of Cameron et al in *Barnard* provides the strongest justification for why a rationality analysis does not suffice.¹²⁹ If the test for valid affirmative action was solely concerned with the connection between ends and means then this would overlook the potential negative impact that affirmative action measures can have. All that would be required is proof that the measure has some

¹²⁴ EEA, s 42.

¹²⁵ EEA, s 15(3) and (4).

¹²⁶ EEA, s 15(1).

¹²⁷ On the distinction between rules, standards, principles and other norms, see CR Sunstein 'Problems with Rules' (1995) 3 *California Law Review* 953. See also Raz (note 1 above) ch 1.

¹²⁸ Albertyn (note 8 above) 730.

¹²⁹ *Barnard* (note 2 above) at paras 94–96.

rational connection to the advancement of a disadvantaged group, ignoring any impact that this measure may have on other groups.

Overlooking the possible negative consequences of affirmative action would be especially problematic where members of historically disadvantaged groups are excluded or actively harmed by these measures. The Constitutional Court's most recent judgment in *Solidarity* now confirms that the *Van Heerden* test will apply even where disadvantaged groups are harmed or excluded. If the impact of the affirmative measure on these groups was not taken into account, this could have serious consequences. To take a stylised example, consider an affirmative action measure that creates bursaries for women to attend university by reallocating all of the budget that was previously assigned for bursaries that were to be awarded to disabled students. A pure test of rationality would not allow any consideration of the impact that this measure may have on the disabled students. It would only be concerned with assessing whether the measure is rationally capable of advancing the legitimate purpose of benefitting women students. There would be no scope for a court to balance the benefits of this measure against the harms to disabled bursary recipients, nor would there be any scope for a court to assess the availability of less restrictive alternatives, such as finding alternative sources of funding to allow bursaries for both groups. Those questions would require a proportionality analysis.

Proportionality assesses whether the benefits outweigh the costs in each case. Proportionality analysis is most commonly associated with the s 36 limitations analysis under the Constitution, but it is not the sole preserve of s 36. Proportionality is applied in areas as diverse as the test for arbitrary deprivations of property,¹³⁰ unfair discrimination,¹³¹ and public participation in law-making,¹³² among others.

Proportionality analysis is also not exclusive to the legal domain. As Rivers argues, proportionality is a tool of practical reasoning that is applied whenever we deliberate about the correct course of action in the face of competing interests and scarce resources:

The doctrine of proportionality is not simply a legal device to assist judges in regulating legislative and executive incursions on rights. It better understood as a rational device for the optimisation of interests.¹³³

The logical structure of proportionality remains the same wherever it is applied, in law or in other decision-making contexts. What differs is the intensity and rigour with which it is applied. As a result, proportionality must be coupled with

¹³⁰ *First National Bank of SA Ltd t/a Wesbank v Commissioner, South African Revenue Service and Another; First National Bank of SA Ltd t/a Wesbank v Minister of Finance* [2002] ZACC 5, 2002 (4) SA 768 (CC), 2002 (7) BCLR 702 (CC) at para 86 (a deprivation of property is arbitrary if it is disproportionate to its purpose, albeit that this is less exacting than a s 36 analysis).

¹³¹ See authors cited at notes 36–37 above.

¹³² *Doctors for Life International v Speaker of the National Assembly and Others* [2006] ZACC 11, 2006 (6) SA 416 (CC), 2006 (12) BCLR 1399 (CC) at para 128 (the test for whether the legislature has acted reasonably in facilitating public participation depends on a balancing of multiple considerations, bearing strong resemblance to a proportionality analysis).

¹³³ Rivers (note 19 above) 207.

a set of principles and other norms to govern the variable intensity of review.¹³⁴ I will say more about these norms in a moment.

Before exploring the ways in which the proportionality analysis can be varied, it is important to understand the logical structure of the proportionality analysis. Where the Constitutional Court has applied a proportionality analysis in other areas of its jurisprudence, it has avoided a 'check-list' approach to proportionality, as is favoured in other jurisdictions.¹³⁵ A check-list approach involves treating the proportionality analysis as a list of requirements that must be ticked off in sequence.¹³⁶ Instead the Constitutional Court opts for what it terms a 'global' judgment on proportionality, meaning that it reaches an all-things-considered judgment on where the balance of interests lies.¹³⁷

While the Court does not approach proportionality in a structured way, the common features of proportionality analysis are reflected in its reasoning.¹³⁸ These are the questions of whether the law or action has a legitimate purpose (*legitimacy*), whether it is capable of achieving that purpose (*suitability*), whether there are less restrictive means available to achieve that purpose (*necessity*), and, on balance, whether the attainment of that purpose outweighs the costs (*balancing*).¹³⁹

The three requirements of the *Van Heerden* test incorporate these elements of the proportionality analysis. The first two requirements reflect the dimensions of legitimacy and suitability. The last requirement involves balancing the competing interests. This balancing would necessarily include some assessment of necessity: whether there are other, less restrictive means to achieve the purpose of protecting or advancing the disadvantaged group. If an affirmative action measure could be achieved in other ways that are less exclusionary or harmful to others, then this is a factor to weigh in the balance.¹⁴⁰

On first glance, the majority judgment in *Van Heerden* appeared to rule out any consideration of necessity. This is reflected in the following passage:

The provisions of s 9(2) do not prescribe ... a necessity test because remedial measures must be constructed to protect or advance a disadvantaged group. They are not predicated on a necessity or purpose to prejudice or penalise others, and so require supporters of the

¹³⁴ Ibid 182.

¹³⁵ On structured proportionality analysis, see generally A Barak *Proportionality: Constitutional Rights and Their Limits* (2012) chs 9–12.

¹³⁶ See, eg, the 'Oakes test' for justifiable limitations under s 1 of the Canadian Charter, developed in *R v Oakes* [1986] 1 SCR 103.

¹³⁷ *Manamela* (note 25 above) at para 32. See N Petersen 'Proportionality and the Incommensurability Challenge in the Jurisprudence of the South African Constitutional Court' (2014) 30 *South African Journal on Human Rights* 405, 407.

¹³⁸ Petersen (note 137 above). While proportionality analyses share common features, courts in different jurisdictions adopt different approaches to this analysis. See further D Grimm 'Proportionality in Canadian and German Constitutional Jurisprudence' (2007) 57 *University of Toronto Law Journal* 383.

¹³⁹ For deeper analysis of these components of the proportionality analysis, see Barak (note 135 above) chs 6, 9–12.

¹⁴⁰ For further examination of this necessity component of the proportionality analysis, and the different forms that necessity analysis can take, see D Bilchitz 'Necessity and Proportionality: Towards a Balanced Approach?' in L Lazarus, C McCrudden & N Bowles (eds) *Reasoning Rights* (2014); Petersen (note 137 above).

measure to establish that there is no less onerous way in which the remedial objective may be achieved.¹⁴¹

The point Moseneke J made in this passage is that the validity of affirmative action does not require proof that there is no other way to protect or advance the disadvantaged group. In doing so, Moseneke J does not claim that consideration of less restrictive alternatives is irrelevant to the enquiry. He simply emphasises that the availability of less restrictive alternatives is not determinative of the validity of an affirmative action measure. In doing so, he is rejecting a test of 'strict' necessity.¹⁴² However, Moseneke J should not be interpreted as dispensing with any consideration of less restrictive alternatives. The availability of alternatives will not be conclusive, but it is a relevant consideration that must weigh in the balance.

While necessity forms a part of this test, courts will generally approach this enquiry with some degree of deference to the initial decision-maker. As Sachs J cautioned in *Van Heerden*:

Courts must be reluctant to interfere with such measures, and exercise due restraint when tempted to interpose themselves as arbiters as to whether the measure could have been proceeded with in a better or less onerous way.¹⁴³

As with all other elements of the *Van Heerden* test, the intensity with which a court evaluates the existence of less restrictive alternatives will depend on the context. I now turn to discuss this variable intensity of review in more detail.

C Taming the Variable Intensity of Review

By affirming that the *Van Heerden* test applies under the EEA and that the test involves a proportionality analysis, the Constitutional Court would go a long way in providing further clarity and guidance for lower courts. As *Barnard* shows, further guidance is also urgently needed on the questions of when and how this intensity of review may be varied in applying the *Van Heerden* test.

Variability in the intensity of review is necessary and unavoidable. The complexities of affirmative action, the different interests at play, and the different contexts in which affirmative action may be applied require some flexibility in the test. In particular, the intensity of review will need to be increased significantly where members of historically disadvantaged groups challenge affirmative action measures that exclude them or compound their disadvantage.

Two things are needed to bring greater guidance and accountability to this variable intensity of review. First, the Court must be conscious and open about the different ways in which the *Van Heerden* test can be applied with variable intensity. Second, it is necessary for the Court to articulate and apply a set of principles and other norms to justify why it has chosen a particular intensity of review in each case.

¹⁴¹ *Van Heerden* (note 3 above) at para 43.

¹⁴² See further Bilchitz (note 140 above).

¹⁴³ Ibid 152.

1 *The Points of Variability*

The three stages of the *Van Heerden* test each allow for variable intensity in their application. The first stage requirement – that historically disadvantaged groups should be in the ‘overwhelming majority’ of beneficiaries of an affirmative action measure – can be applied more or less strictly depending on the circumstances. This was evident in the disagreement between the majority judgment and the separate concurrences in *Van Heerden*. The evidence showed that 79 per cent of the beneficiaries of the affirmative action measure were black. Moseneke J, writing for the majority, concluded that this was an overwhelming majority. Mokgoro J and Ngcobo J held that it was not. What was missing from these decisions was any justification for these different degrees of strictness in applying this requirement.

The second stage requirement – a reasonable likelihood that the affirmative action measure will benefit the disadvantaged group – is also capable of different degrees of scrutiny. In *Van Heerden*, Moseneke J stressed that this requirement does not necessarily require evidence showing that affirmative action actually produces benefits. As discussed above, this evidence will often be hard to come by, given that the long-term effects of affirmative action are often unknowable or hard to predict. Nevertheless, it is notionally possible that courts may vary this requirement depending on the case. If a court is confronted with an affirmative action measure that has a proven negative impact on vulnerable groups it would, in all likelihood, not be satisfied by vague and unsubstantiated predictions about the long-term benefits of the programme.

This second stage requirement is likely to be applied more strictly where the notional beneficiaries of an affirmative action programme are also its victims. Some affirmative action programmes may impose burdens and disadvantages on those they are intended to advance, in a way that entrenches patterns of disadvantage rather than removing them. The recent controversy over the UThukela ‘maidens’ bursary programme is a case in point.¹⁴⁴ A municipality provided bursaries to young women for their tertiary education, on the condition that they remain virgins and subject themselves to invasive virginity testing. This was arguably an affirmative action programme, aimed at providing predominantly black, rural women with educational opportunities. However, the conditions attached to these bursaries reinforced the stigma and double-standards surrounding women’s sexuality, as well as violating the privacy and bodily integrity of its beneficiaries. Measures like this are burdens disguised as benefits. Courts would and should subject the alleged benefits of such measures to more intense scrutiny.

The third requirement of the *Van Heerden* test is perhaps most open to variation. As was just explained, the proportionality analysis has numerous components. Each of these components can be applied more or less stringently in assessing where the balance of interests lies. In particular, a court’s willingness to interrogate the availability of less restrictive alternatives is likely to vary from case to case. In *Barnard* and *Van Heerden*, the Court was reluctant to interrogate alternatives. By contrast, if a court was confronted with a programme such as the UThukela

¹⁴⁴ See further M Smuts ‘State Sponsored Virginity: South Africa’s Maidens’ Bursary Scheme’ *Oxford Human Rights Hub Blog*, available at <http://ohrh.law.ox.ac.uk/state-sponsored-virginity-south-africas-maidens-bursary-scheme/>.

maidens' bursary then the availability of alternative means would hopefully weigh heavily in the analysis. These bursaries could still be offered to young women without the sexist and invasive conditions attached. The need to protect against harmful gender discrimination disguised as a benefit would require more intense scrutiny of these measures.

Another way in which this test can be applied more or less stringently is by adjusting the burden of proof. As indicated above, the burden of proof can be broken down into two elements: the burden of 'non-persuasion' and the evidential and argumentative burden.¹⁴⁵

The burden of non-persuasion entails that where the probabilities are evenly balanced or uncertain, the party bearing the burden will lose.¹⁴⁶ It is clear from the Court's decisions in *Van Heerden* and *Barnard* that the claimant in affirmative action cases will generally bear this burden. However, there is nothing preventing a court from placing this burden on the opposing party in appropriate cases.

The evidential and argumentative burden involves the responsibility to adduce evidence and argument on a point. As the Court recognised in *Mobunram v NDPP*,¹⁴⁷ the evidential and argumentative burden may be shifted, particularly where certain information is 'peculiarly within the knowledge' of one of the parties.

A court also has a choice to come to the assistance of one of the parties where the evidence and reasons provided are not sufficient to discharge the burden. This may involve the court taking judicial notice of facts not on the record or delving into the record to find evidence that supports or contradicts claims made by the decision-maker. It can also involve a court raising points of law *mero motu* either in support or against the affirmative action measure. These are all critical choices. Cameron et al and Van der Westhuizen J hinted that if *Barnard* had staked her claim on gender discrimination then the Court may have been more inclined to take a proactive role in finding evidence and arguments in her favour. In future cases, these choices may make all the difference to affirmative action measures, particularly where disadvantaged groups attack the under-inclusiveness or heavy-handedness of such measures.

There are many other subtle choices that may be made in adjusting the intensity of review in applying the *Van Heerden* test. These choices involve small nuances in the way that rules of procedure, evidence and doctrine are applied. Each of these choices has significance for the outcome of the case and ought to be guided by deeper principles. I now turn to explain these principles.

¹⁴⁵ Barak (note 137) 437; PJ Schwikkard 'The Standard and Burden of Proof and Evidential Duties in Civil Trials' in PJ Schwikkard & others (eds) *Principles of Evidence* (4th Edition, 2015) 616–617. These principles were articulated in *Pillay v Krishna* 1946 AD 946 at 952–953; *South Cape Corporation (Pty) Ltd v Engineering Management Services (Pty) Ltd* 1977 (3) SA 534 (A) at 548A–C.

¹⁴⁶ Chan (note 51 above) 47.

¹⁴⁷ *Mobunram and Another v National Director of Public Prosecutions and Another* [2007] ZACC 4, 2007 (4) SA 222 (CC), 2007 (6) BCLR 575 (CC) at para 75 (Court considered whether the forfeiture provisions under the Prevention of Organised Crime Act 121 of 1998 had been applied correctly).

2 Identifying the Principles

Variable intensity of review is nothing new to South African constitutional law. Commentators have often noted the unexplained variations in the intensity with which the Constitutional Court reviews state action in other areas of its jurisprudence.¹⁴⁸ As a result, there have been many calls for the Court to adopt a set of norms to govern this variable intensity of review; this could be termed a doctrine of deference. This doctrine aims to identify when it is appropriate for courts to defer to the state by adjusting the intensity of review.¹⁴⁹

Commentators have proposed a broad range of principles that should form part of this doctrine. In doing so, they have drawn on the more advanced UK and Canadian jurisprudence on deference.¹⁵⁰ The principles that emerge from these contributions can be grouped under three broad headings: the *interests at stake*, relative *democratic legitimacy*, and relative *institutional competence*.¹⁵¹

In affirmative action cases, the first set of principles, the *interests at stake*, will likely carry substantial weight in deciding how to vary the intensity of review. These principles operate at two levels. First, there is an appreciation of the importance and need for affirmative action measures in general, as affirmed by s 9(2) of the Constitution. This broad constitutional preference will generally tilt the scales in favour of affirmative action measures. Second, there is an appreciation of the specific interests at stake in each case. Where an affirmative action programme with proven benefits to a disfavoured group is under challenge, a court should be more circumspect in reviewing the justifications for this programme, requiring a lower intensity of review, all other things remaining equal. However, where there is an indication that the affirmative action programme has a severe impact, particularly on a historically disadvantaged group, then a court will be justified in making use of the flexibility in the *Van Heerden* test to scrutinise the affirmative action more closely. An intersectional approach is needed here that acknowledges the ways in which different forms of advantage and disadvantage connect and

¹⁴⁸ In the context of legality review, see authors cited at note 28 above. In the context of socio-economic rights, see K McLean *Constitutional Deference, Courts and Socio-Economic Rights in South Africa* (2009); A Pillay 'Economic and Social Rights Adjudication: Developing Principles of Judicial Restraint in South Africa and the United Kingdom' [2013] *Public Law* 599. For further analysis of the Constitutional Court's fluctuating intensity of review in other areas of its case law, see T Roux *The Politics of Principle: The First South African Constitutional Court, 1995-2005* (2013) chs 6–10.

¹⁴⁹ As discussed above, deference and intensity of review are generally inversely related. The greater the degree of deference shown by the court, the lower the intensity of review.

¹⁵⁰ See McLean (note 148 above) ch 1. On the UK's deference jurisprudence, see, eg, Rivers (note 19 above); Kavanagh 'Deference or Defiance?' (note 1 above); J King *Judging Social Rights* (2012). On the Canadian jurisprudence, see, eg, S Choudhry 'So What Is the Real Legacy of *Oakes*? Two Decades of Proportionality Analysis under the Canadian Charter's Section 1' (2006) 34 *Supreme Court Law Review* 501; R Dixon 'The Supreme Court of Canada, Charter Dialogue and Deference' (2009) 47 *Osgoode Hall Law Journal* 235.

¹⁵¹ McLean (note 148 above) chs 1, 2; A Price 'The Content and Justification of Rationality Review' in S Woolman & D Bilchitz (eds) *Is This Seat Taken? Conversations at the Bar, the Bench and the Academy about the South African Constitution* (2012) 50; Hoexter, *Administrative Law* (note 24 above) 151–155; Pillay (note 150 above) 600. This is by no means an exhaustive list, as other principles may also be relevant, such as the availability of non-judicial accountability mechanisms, polycentricity, comity and so on.

overlap.¹⁵² Black women, for example, experience the combined effects of racism and sexism, which create deeper and unique forms of disadvantage which are not shared by black men or white women. In many cases, these intersecting patterns of disadvantage will make stronger demands for redress. For instance, an affirmative action measure that harms or excludes black women would, in general, require more intense scrutiny than one that excludes white women.

Second, concern for *democratic legitimacy* requires courts to recognise that there are some matters on which unelected judges should be cautious about second-guessing the decisions of representative bodies. For example, an employment equity plan that has been developed in a bargaining council, involving representatives of trade unions and employers, will generally be reflective of the particular needs and interests of participants in the industry. A court should be cautious in second-guessing the delicate compromises that have been struck in these forums. However, in some cases democratic legitimacy may necessitate closer scrutiny, particularly where processes of deliberation have failed or have systemically ignored the interests of marginalised groups.¹⁵³ Furthermore, as the Court has stressed, democracy requires the development of a culture of justification, a culture that can be promoted by more searching scrutiny in appropriate cases.¹⁵⁴

Finally, considerations of *relative institutional competence* require courts to recognise the limits of their expertise and to be careful not to second-guess decisions reached by more competent bodies equipped with greater capacity, experience, and information. This consideration can, in many cases, also provide a reason for more intense scrutiny where courts are better equipped to decide matters than other branches of state.¹⁵⁵ The court will also be well-placed to scrutinise measures more closely where the other organs of state have failed to act or have displayed clear incompetence.¹⁵⁶

These principles are flexible and open-ended and must be weighed up in each case. What is needed is for the Court to justify openly its chosen intensity of review in each case with reference to these principles. This is not only necessary to ensure greater openness and accountability in these decisions. It is also required to give lower courts guidance in how to navigate these issues.

¹⁵² See the judgment of Cameron et al in *Barnard* (note 2 above) at paras 152–155 on the need for an intersectional approach to affirmative action. In its unfair discrimination jurisprudence, the Constitutional Court has consistently recognised the need for an intersectional approach. See further *National Coalition for Gay and Lesbian Equality and Another v Minister of Justice and Others* [1998] ZACC 15, 1999 (1) SA 6 (CC), 1998 (12) BCLR 1517 (CC) at para 113; *Bhe and Others v Magistrate, Khayelitsha and Others* [2004] ZACC 17, 2005 (1) SA 580 (CC), 2005 (1) BCLR 1 (CC) and *Hassam v Jacobs NO and Others* [2009] ZACC 19, 2009 (5) SA 572 (CC), 2009 (11) BCLR 1148 (CC).

¹⁵³ See further JH Ely *Democracy and Distrust: A Theory of Judicial Review* (1980) chs 4–6; King (note 150 above) ch 6.

¹⁵⁴ E Murecinik 'A Bridge to Where? Introducing the Interim Bill of Rights' (1994) 10 *South African Journal on Human Rights* 31, 32; McLean (note 148 above) 88, 211; S Fredman 'Adjudication as Accountability' in N Bamforth & P Leyland (eds) *Accountability in the Contemporary Constitution* (2013) 106, 108.

¹⁵⁵ On the intricacies of claims of competence, see further King (note 150 above) chs 7–9, who more accurately subdivides the broad question of competence into issues of polycentricity, expertise, and flexibility.

¹⁵⁶ See further Chan (note 18 above) who argues that the government's claims to have greater competence should be proved with evidence, rather than being accepted on face value.

3 *Respecting the Baseline*

A fully developed doctrine of deference not only contains the principles just outlined, but also includes more determinate norms. These norms set a basic minimum intensity of review. In her writing on proportionality analysis, Cora Chan has labelled these norms the ‘baseline intensity’ of review.¹⁵⁷ A court can scrutinise laws and actions more intensely than the baseline requires, but it cannot sink below this basic minimum intensity. Beyond the baseline, the principles of deference provide guidance on when and how to vary the intensity of review. A doctrine of deference for affirmative action cases should therefore include both a baseline intensity of review and a set of principles to govern variability beyond the baseline.

It is beyond the scope of this article to provide a full account of the baseline requirements that apply in affirmative action cases. However, a few baseline norms are readily apparent. The most essential is that the *Van Heerden* test requires a proportionality analysis. For the reasons I have explained above, it would not be permissible for a Court to vary this standard of review so far that it merely applies a rationality analysis in disguise. The EEA adds further baseline requirements in assessing affirmative action, such as a prohibition on quotas and absolute job reservations. These are minimal, baseline requirements that a court cannot avoid. As explained above, the mere fact that an affirmative action measure is not a quota does not mean that it is valid. It may still fall foul of the *Van Heerden* test.

D Applying the Doctrine of Deference to *Barnard*

It is helpful to analyse how the doctrine of deference could have been applied in *Barnard* in justifying the chosen intensity of review. The majority judgment and the concurrences were clearly motivated by these considerations, but there was no real attempt to use these considerations to explain and justify why the Court’s chosen intensity of review was appropriate.

As indicated earlier, the most important question for all of the judgments in *Barnard* was how to allocate the burden of proof, including the burden of non-persuasion and the evidential and argumentative burden. As I previously explained, in allocating the burden of non-persuasion, the Court is deciding on who should bear the risk of uncertainty about the merits of the affirmative action programme. In affirmative action cases, this burden is generally placed on the persons challenging the affirmative action measure, but it may be shifted in appropriate circumstances. There is no strict, baseline norm requiring that this burden must always fall on one party. This is in contrast with the test for unfair discrimination, as the presumption of unfairness creates a firm baseline norm that the party accused of unfair discrimination will bear the burden of proving that discrimination is fair. In the absence of a baseline norm in affirmative action cases, the burden of non-persuasion will largely be determined by the three principles of deference identified above.

¹⁵⁷ Chan (note 1 above). On the baseline intensity of review under the *Harksen* test for unfair discrimination, see McConnachie (note 37 above).

The first principle, the interests at stake, provided a strong justification for placing the risk of uncertainty on Barnard. At the general level, affirmative action measures must be seen as an important component of achieving equality, in line with s 9(2) of the Constitution. The drafters of the Constitution expressed a general policy preference for affirmative action which courts cannot ignore or dismiss. Principles of democratic legitimacy therefore require courts to avoid second-guessing that general preference.

While the Constitution expresses a general preference for affirmative action, that does not mean that burden of non-persuasion will always be on the claimant. If the claimant can show that the affirmative action measure or its implementation imperils important interests, then the Court may be inclined to shift the risk of non-persuasion to the opposing party. Barnard failed to make out such a case. Had she put up some credible evidence that the implementation of the employment equity plan had a negative impact on gender equality or on service delivery, then the Court may have been more inclined to shift the burden to the National Commissioner. However, in the absence of any compelling interests at stake, Barnard rightly bore the risk of uncertainty.

Issues of institutional competence also have a bearing on this burden. A court will generally assume that senior decision-makers such as the National Commissioner are skilled and competent professionals. Absent any proof to the contrary, they are rightly slow to intervene. As the judgments indicated, the National Commissioner's failure to give evidence and further justification for the decision did complicate matters. Competence is, after all, best demonstrated by the quality of the justification provided for the decision. Nevertheless, in the absence of evidence to the contrary, the National Commissioner's competence was not called into question.

The National Commissioner certainly had an evidential and argumentative burden to justify and expand upon his cursory reasons for refusing to promote Barnard. As Cameron et al noted, accountability requires public officials to be open about their reasoning, particularly where silence can engender bitterness and distrust.¹⁵⁸ However, the fact that the National Commissioner did not discharge his evidential and argumentative burden does not necessarily impact upon the result. Once it is accepted that Barnard bore the risk of non-persuasion, any paucity of evidence on the record was to her disadvantage.

There is some uneasiness about letting the National Commissioner off the hook for his failure to provide a full justification for his actions. However, this uneasiness can be resolved if we acknowledge two things. First, the variable intensity of review means that the National Commissioner would not be let off so lightly if the circumstances were different. If Barnard had based her case on gender equality, then the burden of non-persuasion may well have been placed on the National Commissioner. Similarly, if it was shown that the National Commissioner's implementation of the employment equity plan excluded other designated groups, such as black women or disabled persons, then the Court would have been more likely to shift this burden. Second, it must also be recognised that the burden of non-persuasion is not a difficult burden to bear when the

¹⁵⁸ *Barnard* (note 2 above) at paras 110–111.

justification presented by the other side is scant or non-existent. It would not have been impossible for Barnard to discharge her burden, had a proper case been made out.

In sum, the principles outlined above can be used to offer a persuasive justification for the Court's chosen intensity of review in *Barnard*. Justifications like this are needed, not only for the sake of transparency, but also for the sake of offering guidance to lower courts in navigating these difficult issues.

VII CONCLUSION

Barnard was an opportunity to accomplish three essential tasks. First, clarity was needed on the application of the *Van Heerden* test to affirmative action measures under the EEA. That task was partially achieved when the Court confirmed that the *Van Heerden* test applies to affirmative action measures. However, the uncertainty over whether this test applies to the implementation of these measures leaves a critical gap in the law.

Second, the Court needed to clarify the appropriate standard of review to be applied. In this respect, *Barnard* is a step backwards as the divergent approaches are likely to produce greater uncertainty.

Third, the Court needed to clarify when and how it is appropriate to vary the intensity of review in scrutinising affirmative action measures under the *Van Heerden* test. As I have argued, flexibility in the application of the *Van Heerden* test is important, provided that clear and principled reasons are provided for applying the test more or less stringently. This flexibility is apparent in the way that the court modifies the standard of review, adjusts the evidential and argumentative burden, and takes it upon itself to consider evidence and arguments not squarely raised by the parties. Future cases will likely turn on these decisions. Indeed, in *Barnard*, the burden of proof proved to be the decisive matter in all judgments.

I have argued for three developments in the way in which the Court scrutinises affirmative action. First, the *Van Heerden* test should be applied under the EEA and other specialist legislation in considering both affirmative action measures *and* the implementation of these measures. Second, the Court must clarify that the *Van Heerden* test involves a proportionality analysis. Finally, and most significantly, the Court needs to develop and apply a doctrine of deference to justify variations in the intensity with which it applies the *Van Heerden* test.

This doctrine of deference has two parts. First, there are the baseline requirements, the non-negotiable minimum level of scrutiny that must be applied. Second, and more significantly in this area, there are the principles that must be weighed against each other to justify the intensity of review that is applied beyond the baseline.

Determining the appropriate intensity of review is ultimately a value-laden exercise. While these decisions manifest in the application of doctrines and formulae, they are driven by a deeper understanding of the appropriate and legitimate institutional role of the courts in adjudicating challenges to affirmative action. Unless a more transparent, principled approach is adopted here, the Court's affirmative action jurisprudence will remain obscure and uncertain.

The *Zimbabwe Torture*
Docket Case

South Africa's Competing Obligations in Relation to International Crimes

Lilian Chenwi*
Franziska Sucker†

I INTRODUCTION

The need to end impunity for international crimes is largely recognised by the international community.¹ Accordingly, significant obligations have been imposed, by both conventional (treaty) and customary international law, on states to prosecute certain international crimes.² In addition to the duty to prosecute, conventional international law also imposes the duty to investigate allegations of international crimes³ as well as the duty to cooperate in the investigation and prosecution of international crimes.⁴ Conventional law has in fact accentuated a duty upon states to exercise jurisdiction over international crimes.⁵

South Africa's domestic legal order, at least in theory, is receptive to these obligations, including having in place a legal framework for the enforcement of international criminal law within the country. This is evidenced from its ratification and implementation of treaties that impose these obligations on South Africa in relation to certain international crimes.⁶

* Professor, School of Law, University of the Witwatersrand.

† Senior Lecturer, School of Law, University of the Witwatersrand.

¹ This is spelt out in various documents. For example, the determination to end impunity for 'the most serious crimes of concern to the international community' is stated in the preamble to the Rome Statute of the International Criminal Court, UN Doc A/CONF183/9 (17 July 1998), reprinted in (1998) 37 *ILM* 1002 ('Rome Statute'). African states have also reiterated their condemnation and rejection of impunity and commitment to fight impunity in the preamble to the Protocol on Amendments to the Statute of the African Court of Justice and Human Rights (2014). In simple terms, an international crime is 'an offence which is created by international law' (R Cryer, H Friman, D Robinson & E Wilmshurst *An Introduction to International Criminal Law and Procedure* (2nd Edition, 2010) 8) or, put differently, 'any act entailing the criminal liability of the perpetrator and emanating from treaty or custom' (I Bantekas *International Criminal Law* (4th Edition, 2010) 8).

² See B van Schaack & RC Slye *International Criminal Law and its Enforcement: Cases and Materials* (2nd Edition, 2010) 21.

³ See, eg, the Convention against Torture and Other Cruel, Inhuman or Degrading Treatment or Punishment (1984) art 11 (Requires states parties to promptly and impartially investigate allegations of torture).

⁴ Article 86 of the Rome Statute (Places a general obligation on states parties to 'cooperate fully with the [International Criminal] Court in its investigation and prosecution of crimes within the jurisdiction of the Court').

⁵ See, eg, preamble to the Rome Statute (Recalls 'that it is the duty of every State to exercise its criminal jurisdiction over those responsible for international crimes').

⁶ The South African government ratified the Rome Statute on 27 November 2000 and implemented it in 2002. Implementation of the Rome Statute Act 27 of 2002 ('ICC Act'). See also the Implementation of the Geneva Conventions Act 8 of 2012 ('GCA') and the Prevention and Combating of Torture of Persons Act 13 of 2013 ('Torture Act').

The country's relationship with international criminal law in practice is, however, controversial and has been described as 'complex' and 'schizophrenic'.⁷ South Africa has taken contradictory positions on issues relating to international criminal law and justice. For example, the South African government participated in the drafting of the Rome Statute of the International Criminal Court (Rome Statute)⁸ and supported the referral of Libya to the International Criminal Court (ICC). But four months after the referral, it joined other states in requesting the deferral of the case.⁹ Also, despite having what has been described as 'the most progressive legislation for the prosecution of international crimes by its courts', the government has been reluctant to effectively implement the legislation.¹⁰ Most controversially, on 19 October 2016, the government submitted an Instrument of Withdrawal from the Rome Statute to the United Nations (UN) Secretary General.¹¹ Following a High Court decision,¹² it withdrew the Instrument of Withdrawal.¹³

The complex nature of South Africa's relationship with the ICC is compounded by the position of the African Union (AU). The AU has opposed the prosecution of international crimes in the exercise of universal jurisdiction by non-African states, has argued that the ICC is selective or biased, and has decided that AU member states should not cooperate with the ICC in the execution of the arrest warrants issued by the ICC against President Omar al-Bashir of The Sudan (al-Bashir) and the late Colonel Qadhafi of Libya.¹⁴ The AU's position places

⁷ See, eg, C Gevers 'International Criminal Law in South Africa' in E de Wet, H Hestermeyer & R Wolfrum (eds) *The Implementation of International Law in Germany and South Africa* (2015) 403–404.

⁸ Rome Statute (note 1 above).

⁹ See Gevers (note 7 above) at 403.

¹⁰ Ibid at 404.

¹¹ See 'Declaratory statement by the Republic of South Africa on the Decision to Withdraw from the Rome Statute of the International Criminal Court' quoted in 'Rome Statute of the International Criminal Court Rome, 17 July 1998–South Africa: Withdrawal', C.N.786.2016.TREATIES-XVIII.10 (Depositary Notification) (25 October 2016), available at <https://treaties.un.org/doc/Publication/CN/2016/CN.786.2016-Eng.pdf> ('Instrument of Withdrawal'). In accordance with art 127(1) of the Rome Statute, South Africa's 'withdrawal shall take effect one year after the date of receipt of the notification, unless the notification specifies a later date'. On the constitutionality of this controversial action, see part VII.

¹² *Democratic Alliance v Minister of International Relations and Cooperation and Others (Council for the Advancement of the South African Constitution Intervening)* [2017] ZAGPPHC 53, 2017 (3) SA 212 (GP), [2017] 2 All SA 123 (GP).

¹³ See 'South Africa revokes ICC withdrawal', Mail & Guardian (8 March 2017), available at <https://mg.co.za/article/2017-03-08-south-africa-revokes-icc-withdrawal>.

¹⁴ See *Decision on the Meeting of African States Parties to the Rome Statute of the International Criminal Court (ICC)* Doc Assembly/AU/13(XIII) (1–3 July 2009) at para 10 and *Decision on the Implementation of the Assembly Decisions on the International Criminal Court*, Doc Ex.Cl/670(XIX), contained in Assembly/AU/Dec.366(XVII)(30 June–1 July 2011) at para 6. For further reading on the relationship between the ICC and AU/Africa and the AU's decision to not cooperate with the ICC, see, eg, T Murithi 'The African Union and the International Criminal Court: An Embattled Relationship' (2013) 8 *Institute for Justice and Reconciliation Policy Brief* 1–9; M du Plessis, T Maluwa & A O'Reilly *Africa and the International Criminal Court* (2013), available at https://www.chathamhouse.org/sites/files/chathamhouse/public/Research/International%20Law/0713pp_iccafrica.pdf; and International Commission of Jurists, *Kenya Reflections on the African Union ICC Relationship* (2014), available at <http://www.icj-kenya.org/dmdocuments/books/reflections%20on%20the%20african%20union%20icc%20relationship.pdf>. It is worth noting that the tension between the African states and the ICC in relation to al-Bashir has resulted in the ICC 'threatening to refer the matter to the UN Security Council'. See 'ICC Threatens

African states that are parties to the Rome Statute in a difficult position. They have obligations towards the AU that they must comply with, or risk sanctions; while at the same time they have obligations under the Rome Statute to cooperate with the ICC in the investigation and prosecution of international crimes.¹⁵

The Kenyan government, for example, found itself in exactly this position when President al-Bashir, against whom two arrest warrants have been issued by the ICC,¹⁶ visited its country in 2010. The government did not arrest him on the basis that it had to balance its obligations towards the AU with those towards the ICC. Kenya's preference for compliance with its AU obligations resulted in non-compliance with its obligations under the Rome Statute: a position that was endorsed by the AU.¹⁷

The South African government was in a similar position when al-Bashir visited the country in 2015. The government did not arrest him, primarily on the basis of al-Bashir's incumbent Head of State immunity, and he subsequently left the country.¹⁸ This was done in the face of two interim High Court orders prohibiting al-Bashir from leaving South Africa and directing the government 'to take all necessary steps to prevent him from doing so';¹⁹ and one further High Court order of the Full Bench stating that the government's failure 'to take steps to arrest and/or detain' al-Bashir was in contravention of the Constitution and thus invalid. The second order also required the government forthwith 'to take all reasonable steps'²⁰ to arrest and detain al-Bashir, pending a formal request from

to Refer al-Bashir case to UN' Legalbrief (22 February 2017), available at <http://legalbrief.co.za/diary/legalbrief-africa-new/story/icc-threatens-to-refer-al-bashir-case-to-un/pdf>.

¹⁵ These conflicts are discussed in Part C.3. See also M du Plessis & C Gevers 'Balancing Competing Obligations: The Rome Statute and AU Decisions' (2011) *ISS Paper* 225, 2–3.

¹⁶ ICC Pre-Trial Chamber I issued the arrest warrants after it considered that 'there are reasonable grounds to believe that Omar al-Bashir is criminally responsible as an indirect perpetrator, or as an indirect co-perpetrator' for crimes against humanity and war crimes (in relation to the first arrest warrant issued on 4 March 2009) and three counts of genocide (in relation to the second arrest warrant issued on 12 July 2010). See, generally, *Prosecutor v Omar Hassan Ahmad al-Bashir (Warrant of Arrest for Omar Hassan Ahmad al-Bashir)* ICC-02/05-01/09-1 (4 March 2009) PT Ch I; and *Prosecutor v Omar Hassan Ahmad al-Bashir* (Second Decision on the Prosecution's Application for a Warrant of Arrest) ICC-02/05-01/09-94 (12 July 2010) PT Ch I.

¹⁷ See du Plessis & Gevers (note 15 above) at 3.

¹⁸ The government has subsequently stated in its Instrument of Withdrawal from the Rome Statute that it 'was faced with the conflicting obligation to arrest President al-Bashir under the Rome Statute, the obligation to the AU to grant immunity in terms of the Host Agreement, and the General Convention on the Privileges and Immunities of the Organization of African Unity of 1965 as well as the obligation under customary international law which recognises the immunity of sitting heads of state'. Instrument of Withdrawal (note 11 above) at 2.

¹⁹ *Minister of Justice and Constitutional Development and Others v Southern Africa Litigation Centre and Others* [2016] ZASCA 17 (SCA), 2016 (4) BCLR 487 (SCA), [2016] 2 All SA 365 (SCA), 2016 (3) SA 317 (SCA) ('al-Bashir') at para 5 (Reproduces the High Court orders).

²⁰ Ibid at para 6. For further details on this see MJ Ventura 'Escape From Johannesburg? Sudanese President Al-Bashir Visits South Africa, and the Implicit Removal of Head of State Immunity by the UN Security Council in light of *Al-Jedda*' (2016) 13(5) *Journal of International Criminal Justice* 995, esp 1001–1005.

the ICC for his surrender.²¹ That decision was subsequently upheld, although for different reasons, by the Supreme Court of Appeal (*al-Bashir*).²² A previous case – *SALC* – on South Africa’s obligations ‘to investigate crimes against humanity’²³ also illustrates the complex nature, in practice, of South Africa’s relationship with international criminal justice.

SALC and *al-Bashir* concern South Africa’s obligations in relation to the investigation and cooperation in the prosecution of international crimes, and in relation to immunity. In the light of these cases, we consider South Africa’s competing international and domestic obligations regarding international crimes. We identify the specific obligations that South Africa has under both international and domestic (South African) law, establish the obligations that are in possible conflict from the perspective of international as well as domestic law, and discuss approaches to addressing the possible conflicts.

In Part II, we briefly introduce the issues in *SALC* and *al-Bashir*. We then set out, in Part III, the basis for the South African government’s exercise of jurisdiction over international crimes by looking at both international and domestic law. Thereafter, we outline South Africa’s obligations in relation to the investigation and prosecution of international crimes as stipulated in both international and domestic law, and elaborate on conflicts between obligations to (not) cooperate stemming from different levels and their potential resolution (in Part IV). After establishing the relevant principles and obligations relating to immunity and international crimes in Part V, we consider the conflict between South Africa’s cooperation obligations and obligations relating to immunity and whether and how such a conflict can be resolved (in Part VI).

II BRIEF INTRODUCTION TO *AL-BASHIR* AND *SALC*

A *Al-Bashir*

Al-Bashir, in the High Court, raised the question of South Africa’s obligations in the context of the ICC Act, specifically in relation to the arrest of an incumbent head of state against whom arrest warrants for international crimes had been issued

²¹ On 6 March 2009, Pre-Trial Chamber I of the ICC requested all states parties to the ICC Statute to arrest and surrender the Sudanese President for trial by the ICC. See *Prosecutor v Omar Hassan Ahmad al-Bashir* (Request to All States Parties to the Rome Statute for the Arrest and Surrender of Omar al-Bashir) Case ICC-02/05-01/09-7 (6 March 2009) PT Ch I; and *Prosecutor v Omar Hassan Ahmad al-Bashir* (Supplementary Request to All States Parties to the Rome Statute for the Arrest and Surrender of Omar Hassan Ahmad al-Bashir) Case ICC-02/05-01/09-96 (21 July 2010) PT Ch I.

²² *Al-Bashir* (note 19 above).

²³ *National Commissioner of the South African Police Service v Southern African Human Rights Litigation Centre and Another* [2014] ZACC 30, 2015 (1) SA 315 (CC), 2015 (1) SACR 255 (CC), 2014 (12) BCLR 1428 (CC) (*SALC*) at para 3.

by the ICC.²⁴ The High Court held that immunity, though recognised under customary international law, could not be advanced as a ground for not arresting al-Bashir because that immunity was not applicable in light of the Implementation of the Rome Statute Act 27 of 2002 (ICC Act).²⁵ It also held that South Africa has domestic and international obligations in relation to the arrest of al-Bashir, which it is 'bound to comply with'.²⁶ Otherwise, it would result in the collapse of the 'democratic edifice' and the rule of law.²⁷

On appeal by the government, the issues before the Supreme Court of Appeal (SCA) included whether al-Bashir enjoyed immunity from arrest and whether such immunity had been waived.²⁸ We address the SCA's ruling in more detail below. In a nutshell, the SCA held that the government's failure to arrest and surrender al-Bashir was unlawful and contravened South Africa's obligations under the Rome Statute and the ICC Act.²⁹ The government again appealed to the Constitutional Court (CC)³⁰ but later announced that the appeal would be withdrawn, following its decision to withdraw from the Rome Statute.³¹

B *SALC*

The question in *SALC* related to 'the extent to which the South African Police Service (SAPS) has a duty to investigate allegations of torture [which constituted crimes against humanity] committed in Zimbabwe by and against Zimbabwean nationals'.³² The Court had to determine South Africa's obligations – both international and domestic – 'to prevent impunity' and ensure accountability for international crimes committed beyond its borders and by foreign nationals.³³ Taking into consideration the international and domestic obligations, the Court then had to establish whether 'the SAPS has a duty to investigate crimes against humanity committed beyond [SA] borders' and if it does, 'under which circumstances is this duty triggered'.³⁴ At first instance, the High Court held that the decision not to initiate an investigation under the ICC Act was unconstitutional and unlawful; and that immunity and other considerations are not relevant at

²⁴ *Southern Africa Litigation Centre v Minister of Justice and Constitutional Development and Others* [2015] ZAGPPHC 402, 2016 (1) SACR 161 (GP), 2015 (5) SA 1 (GP), [2015] 3 All SA 505 (GP), 2015 (9) BCLR 1108 (GP) at para 1 ('*al-Bashir HC*'). (The court stated the question in the case as follows: '[W]hether a cabinet resolution coupled with a ministerial notice are capable of suspending this country's duty to arrest a head of state against whom the International Criminal Court (ICC) has issued arrest warrants for war crimes, crimes against humanity and genocide').

²⁵ Ibid at paras 28–32.

²⁶ Ibid at para 37.1.

²⁷ Ibid at paras 37.2 and 38.

²⁸ See *al-Bashir* (note 19 above) at para 18.

²⁹ Ibid at para 113.

³⁰ *Minister of Justice and Constitutional Development and Others v Southern African Litigation Centre* CCT 75/16.

³¹ See 'SALC Statement on South Africa's Withdrawal from the ICC and Withdrawal of Bashir Constitutional Court Case' (24 October 2016), available at <http://www.southernafricalitigationcentre.org/2016/10/24/salc-statement-on-south-africas-withdrawal-from-the-icc-and-withdrawal-of-bashir-constitutional-court-case-24-october/>.

³² *SALC* (note 23 above) at para 4.

³³ Ibid.

³⁴ Ibid at para 21.

the investigation stage.³⁵ On appeal, the SCA upheld the High Court's decision, stating that the SAPS has a duty to 'initiate an investigation into the alleged acts of torture' regardless of 'whether or not the alleged perpetrators are present in South Africa'. That obligation stems from the Constitution, the South African Police Service Act³⁶ and the ICC Act.³⁷

On further appeal, the CC acknowledged that the issue was of 'substantial complexity', but still held that the SAPS 'must investigate the complaint'.³⁸ This duty, the Court held, arises from both domestic law and international law and 'must be honoured'.³⁹ The South African government, Majiedt AJ held, 'cannot be seen to be tolerant of impunity for alleged torturers' or 'a safe haven for those who commit crimes against humanity'.⁴⁰

III JURISDICTION OVER INTERNATIONAL CRIMES

The concept of jurisdiction refers to the power or competence of a state under international law to regulate affairs or the conduct of persons.⁴¹ States can exercise jurisdiction in three ways: prescriptive (legislative), adjudicative (judicial), and enforcement (executive).⁴²

There are five points worth highlighting in relation to the forms of jurisdiction. Firstly, as pointed out by the *SALC* Court, adjudicative jurisdiction is not limited to the enforcement of criminalised conduct through prosecutions but includes investigation.⁴³ Secondly, adjudicative jurisdiction is not limited to the actions of domestic courts but extends to the actions of a state's prosecutorial authorities.⁴⁴ Thirdly, enforcement jurisdiction grants states the right, through their law enforcement agencies, to carry out legal processes such as arrest.⁴⁵ Fourthly, while the competence of states in the exercise of jurisdiction, under international law, is traditionally dependent on the existence of connections (territoriality, nationality, passive personality, protection of the state),⁴⁶ jurisdiction can also be exercised

³⁵ *Southern African Litigation Centre and Another v National Director of Public Prosecutions and Others* [2012] ZAGPPHC 61; 2012 (10) BCLR 1089 (GNP); [2012] 3 All SA 198 (GNP) ('*SALC HC*') at paras 31 and 33. See also *SALC* (note 23 above) at para 16 ('[I]nconsistent with the Constitution and South Africa's international law obligations').

³⁶ Act 68 of 1995 ('SAPS Act').

³⁷ *SALC* (note 23 above) at paras 17–18 and 70.

³⁸ *Ibid* at paras 83–84.

³⁹ *Ibid* at para 80.

⁴⁰ *Ibid* at paras 61 and 80.

⁴¹ See J Crawford Brownlie's *Principles of Public International Law* (8th Edition, 2012) 456; and MN Shaw *International Law* (7th Edition, 2014) 469. See also A Mills 'Rethinking Jurisdiction in International Law' (2014) 84 *British Yearbook of International Law* 187, 194 ('[T]he scope of regulatory authority of states in international law').

⁴² For further reading, see R O'Keefe 'Universal jurisdiction: Clarifying the basic concept' (2004) 2 *Journal of International Criminal Justice* 736–737. See also *SALC* (note 23 above) at para 25; Cryer, Friman, Robinson & Wilmshurst (note 1 above) at 44.

⁴³ *SALC* (note 23 above) at para 25 (emphasis added).

⁴⁴ See C Kreß 'Universal Jurisdiction over International Crimes and the *Institut de Droit international*' (2006) 4(3) *Journal of International Criminal Justice* 564, fn 16.

⁴⁵ See Cryer, Friman, Robinson & Wilmshurst (note 1 above) at 44.

⁴⁶ For a detailed discussion of these jurisdictional bases, see *ibid* at 46–50; Shaw (note 41 above) at 474–485; and Bantekas (note 1 above) at 332–344. See also *SALC* (note 23 above) at paras 26–27 (Acknowledged these four bases of jurisdiction).

in some instances without the existence of such jurisdictional links (universal jurisdiction).⁴⁷ Lastly, though the jurisdictional bases above are seen as consistent with international law, the specific obligation on states to exercise jurisdiction on one or any of the above grounds is generally provided for in domestic law.⁴⁸

In this Part, we expand on how jurisdictional questions affect South Africa's competing obligations. We begin with a discussion of universal jurisdiction in international law. Next, we consider jurisdiction under the Rome Statute. We then look at jurisdiction in South African law and, lastly, how the courts applied these different jurisdictional regimes in *SALC* and *al-Bashir*.

A Universal Jurisdiction in International Law

The exercise of universal jurisdiction, which was at the core of *SALC*, is justified on 'the severity of the crime and the undesirable consequences of impunity'.⁴⁹ However, states' understanding and incorporation of the principle varies, creating uncertainties about its definition and legal status.⁵⁰ There have thus been several efforts to define the principle.⁵¹ Based on the lack of traditional jurisdictional connections in its exercise, universal jurisdiction is defined as 'jurisdiction established over a crime without reference to the place of perpetration, the nationality of the suspect or the victim or any other recognized linking point between the crime and the prosecuting State'.⁵² In practice, however, states exercise universal jurisdiction subject to certain prerequisites such as 'the existence of a specific ground for universal jurisdiction, a sufficiently clear definition of the offence and its constitutive elements, and national means of enforcement allowing the national judiciary to exercise their jurisdiction over these crimes'.⁵³

⁴⁷ Discussed in Shaw (note 41 above) at 485–497; Cryer, Friman, Robinson & Wilmschurst (note 1 above) at 50–62; and Bantekas (note 1 above) at 344–349.

⁴⁸ See Shaw (note 41 above) at 474. See also Van Schaack & Slye (note 2 above) at 27.

⁴⁹ Van Schaack & Slye (note 2 above) at 113. The AU has adopted three resolutions in which it stated the need to close the impunity gap that too often permits perpetrators of grave international crimes to escape justice as the rationale for universal jurisdiction in international law. See Decision on the Abuse of the Principle of Universal Jurisdiction (1–3 July 2009) AU Doc Assembly/AU/11(XIII); Decision on the Implementation of the Assembly Decision on the Abuse of the Principle of Universal Jurisdiction (1–3 February 2009) AU Doc Assembly/AU/3(XII); and Decision on the Report of the Commission on the Abuse of the Principle of Universal Jurisdiction AU (30 June–1 July 2008) Doc Assembly/AU/14(XI).

⁵⁰ See dissenting opinion of Van Den Wyngaert J in *Arrest Warrant of 11 April 2000 (Democratic Republic of the Congo v Belgium)* (2002) ICJ Reports 137 at paras 44–45.

⁵¹ See, eg, the *Princeton Principles on Universal Jurisdiction* (2001) principle 1(1), available at <http://hrlibrary.umn.edu/instree/princeton.html> (Defines universal jurisdiction as 'criminal jurisdiction based solely on the nature of the crime, without regard to where the crime was committed, the nationality of the alleged or convicted perpetrator, the nationality of the victim, or any other connection to the state exercising such jurisdiction'). See also the *Madrid-Buenos Aires Principles on Universal Jurisdiction* (2015), available at <http://www.hormantruth.org/ht/sites/default/files/files/universal%20jurisdiction/MADRID%20-%20BUENOS%20AIRES%20PRINCIPLES%20OF%20UNIVERSAL%20JURISDICTION%20%20%20-EN.pdf>.

⁵² Cryer, Friman, Robinson & Wilmschurst (note 1 above) at 50–51.

⁵³ X Philippe 'The Principles of Universal Jurisdiction and Complementarity: How Do the Two Principles Intermesh' (2006) 88(862) *International Review of the Red Cross* 379 (Outlining necessary steps for its exercise).

The *SALC* Court recognised that international law supports the exercise of universal jurisdiction subject to certain principles.⁵⁴ Drawing from legal writing, the court outlined three of such principles:

- (a) 'there should be a substantial and bona fide connection between the subject-matter and the source of the jurisdiction'; (b) 'the principle of non-intervention in the domestic or territorial jurisdiction of other states should be observed'; and (c) 'elements of accommodation, mutuality, and proportionality should be applied'.⁵⁵

States have adopted different approaches to exercising universal jurisdiction. Relying on universal jurisdiction to prosecute or adjudicate is often conditioned on the presence of the accused within the territory of the concerned state, often referred to as conditional universal jurisdiction. Universal jurisdiction can also be exercised *in absentia*, often referred to as absolute universal jurisdiction.⁵⁶ Many states, however, restrict the exercise of universal jurisdiction to prosecute to the former.⁵⁷ The *SALC* Court was therefore of the view that *investigations* that do not breach the principle of non-intervention are not at odds with the exercise of universal jurisdiction.⁵⁸

The basis for the assertion of universal jurisdiction is found in both customary and conventional international law.⁵⁹ Though some treaties contain provisions that allow for its exercise,⁶⁰ and despite the wide acceptance of its application for international crimes such as genocide, crimes against humanity and war crimes (which are crimes in the Rome Statute),⁶¹ the Rome Statute itself does not make reference to, or explicitly require states to exercise, universal jurisdiction.

B Jurisdiction in the Rome Statute

The Rome Statute refers to two of the conventional jurisdictional bases for the ICC to exercise jurisdiction: territoriality (jurisdiction *ratione loci*) and nationality

⁵⁴ See *SALC* (note 23 above) at para 27.

⁵⁵ Ibid at para 28, quoting I Brownlie *Principles of Public International Law* (5th Edition, 1998) 313.

⁵⁶ See L Chenwi 'Universal Jurisdiction and South Africa's Perspective on the Investigation of International Crimes (2014) *South African Law Journal* 27, 32–33.

⁵⁷ Ibid at 33.

⁵⁸ Ibid at 29.

⁵⁹ This is relevant to *SALC* since it has been argued, in relation to the exercise of universal jurisdiction over torture, that 'permissive universal jurisdiction' over the crime of torture may be seen to constitute a customary international law norm. See Association for the Prevention of Torture & Centre for Justice and International Law *Torture in International Law: A Guide to Jurisprudence* (2008) 21 (Permissive universal jurisdiction is also explained as 'meaning that all States have the legal capacity, but not the obligation, to exercise universal jurisdiction over torture').

⁶⁰ See, eg, Convention against Torture and other Cruel, Inhuman and Degrading Treatment or Punishment (1984)(CAT) art 5(2); Hague Convention for the Suppression of Unlawful Seizure of Aircraft (1970) art 4(2); Geneva Convention for the Amelioration of the Condition of the Wounded and Sick in Armed Forces in the Field (1949)(Geneva Convention I) art 49; Geneva Convention for the Amelioration of the Condition of Wounded, Sick and Shipwrecked Members of Armed Forces (1949)(Geneva Convention II) art 50; Geneva Convention Relative to the Treatment of Prisoners of War (1949)(Geneva Convention III) art 129; Geneva Convention relative to the Protection of Civilian Persons in Time of War (1949)(Geneva Convention IV) art 146; and Protocol Additional to the Geneva Conventions of 12 August 1949, and relating to the Protection of Victims of International Armed Conflicts (1977)(Additional Protocol I) arts 11 and 85.

⁶¹ See WA Schabas *An Introduction to the International Criminal Court* (4th Edition, 2011) 64.

(jurisdiction *ratione personae*).⁶² In addition, ICC jurisdiction is limited in two further ways: temporal (jurisdiction *ratione temporis*) ie 'jurisdiction exists only with respect to crimes committed after the entry into force of [the] Statute';⁶³ and in relation to the subject matter (jurisdiction *ratione materiae*) ie jurisdiction exists only over the listed crimes – genocide, crimes against humanity, war crimes and, following a 2010 amendment, the crime of aggression.⁶⁴

The ICC's jurisdiction can be triggered through state party referral,⁶⁵ UN Security Council (UNSC) referral,⁶⁶ the Prosecutor acting *proprio motu*⁶⁷ or an ad hoc declaration by a state that is not a party to the Rome Statute.⁶⁸ In the case of a UNSC referral, state consent is not required and is, arguably, inferred from UN membership and state obligations under the 1945 Charter of the United Nations (UN Charter).⁶⁹ This was the situation in relation to al-Bashir: the UNSC, acting under Chapter VII of the UN Charter, referred the case to the ICC.⁷⁰ Its jurisdiction is generally narrower than the jurisdiction that individual states are entitled to exercise with respect to these crimes. For example, the ICC is only able to exercise jurisdiction when a state is 'unwilling' or 'unable' to prosecute (principle of complementarity).⁷¹

C Jurisdiction in Domestic Law

The South African government ratified the Rome Statute on 27 September 2000. Parliament domesticated the Statute in August 2002 through the ICC Act, the country's first domestic legislation on implementation of international crimes.

Temporal jurisdiction under the ICC Act is limited to crimes committed after the Act came into force.⁷² Moreover, the Act explicitly identifies the various available jurisdictional bases in relation to crimes against humanity, genocide and war crimes. South African courts can exercise jurisdiction on the basis of nationality (ie the accused or the victim is a South African citizen), ordinary

⁶² Rome Statute art 12(2) which reads:

'In the case of article 13, paragraph (a) or (c), the Court may exercise its jurisdiction if one or more of the following States are Parties to this Statute or have accepted the jurisdiction of the Court in accordance with paragraph 3:

(a) 'The State on the territory of which the conduct in question occurred or, if the crime was committed on board a vessel or aircraft, the State of registration of that vessel or aircraft;

(b) 'The State of which the person accused of the crime is a national.'

⁶³ Rome Statute art 11(1).

⁶⁴ See Rome Statute arts 5, 6, 7, 8 and 8*bis*. It should be noted that the crime of aggression before the ICC is subject to a special jurisdictional regime. With regard to a state that is not a party to the Statute, the ICC 'shall not exercise its jurisdiction over the crime of aggression when committed by that State's nationals or on its territory'. Rome Statute art 15 *bis*(5).

⁶⁵ Rome Statute art 13(a).

⁶⁶ Rome Statute art 13(b).

⁶⁷ Rome Statute arts 13(c) and 15.

⁶⁸ Rome Statute art 12(3).

⁶⁹ Charter of the United Nations (1945)(see TS No 993; 3 Bevans 1153; 59 Stat 1031; (1979) YBUN 1043)(UN Charter).

⁷⁰ See UN Security Council Resolution 1593 (31 March 2005), SCOR (Res & Dec) 131, UN Doc S/RES/1593 (2005)(Resolution 1593).

⁷¹ See Rome Statute preamble and arts 1 and 17. The *SALC* Court made South Africa's exercise of 'universal jurisdiction' subject to an 'unwilling and unable' test. *SALC* (note 23 above) at para 61.

⁷² See ICC Act s 5(2).

residence (ie the accused or the victim is ordinarily resident in South Africa), or the presence of the accused in South Africa after the commission of the crime.⁷³ A distinctive feature of the Act is its recognition that crimes committed outside of South Africa are deemed to have been committed in South African territory.⁷⁴ The Act thus goes beyond the traditional grounds of nationality, territoriality, and passive personality to expressly include universal jurisdiction.⁷⁵

An additional legal basis for South African courts' exercise of universal jurisdiction over certain international crimes,⁷⁶ more precisely acts of torture, is found in the Prevention and Combating of Torture of Persons Act (Torture Act).⁷⁷ By permitting universal jurisdiction, it gives effect to South Africa's obligations under the Convention against Torture and Other Cruel, Inhuman or Degrading Treatment or Punishment of 1984 (CAT). The jurisdictional grounds identified in the Torture Act are similar to those in the ICC Act: nationality or ordinary residence of the accused or the victim, and the presence of the accused in South Africa after the commission of the offence.⁷⁸ And like the ICC Act, the Torture Act grants universal jurisdiction over offences committed outside of South Africa that would have constituted an offence in the country, 'regardless of whether or not the act constitutes an offence at the place of its commission'.⁷⁹

A key point to note in both the ICC Act and Torture Act is the procedural limitation in the exercise of jurisdiction to prosecute. The 'consent' of the National Director of Public Prosecutions (NDPP) is required before any prosecution can be instituted under either Act.⁸⁰ The consent requirement explicitly relates to 'prosecution'; it is not applicable to investigation. This distinction is important as the initiation of an investigation, and not a prosecution, was at issue in *SALC*. Under the ICC Act, a decision by the NDPP on whether or not to prosecute must have due regard to South Africa's international obligations and the principle of complementarity.⁸¹ However, a decision not to prosecute does not bar the

⁷³ See ICC Act s 4(3).

⁷⁴ Ibid.

⁷⁵ South African courts have, through the Act, been granted 'relatively expansive jurisdiction' in relation to Rome Statute crimes. See L Stone 'Implementation of the Rome Statute in South Africa' in C Murungu & J Biegon (eds) *Prosecuting International Crimes in Africa* (2011) 311.

⁷⁶ Though not of relevance to *SALC* or *al-Bashir*, it is worth noting that the GCA permits the South African government to exercise universal jurisdiction over grave and other breaches of the Geneva Conventions of 1949 and Protocols of 1977, which are treaties that apply in situations of armed conflict. See GCA arts 6(2) and 7.

⁷⁷ Act 13 of 2013. The Torture Act incorporates the CAT, which the South African government ratified in 1998.

⁷⁸ Torture Act s 6(1). The territory includes the accused's presence in territorial waters or 'on board a ship, vessel, off-shore installation, a fixed platform or aircraft registered or required to be registered' in South Africa.

⁷⁹ Ibid.

⁸⁰ See ICC Act s 5(1); and Torture Act s 6(2). In terms of the Constitution s 179(1) and (2), and National Prosecuting Authority Act 32 of 1998 s 20, the NDPP is 'the head of the prosecuting authority' with 'the power to institute criminal proceedings on behalf of the state, and to carry out any necessary functions incidental to instituting criminal proceedings'.

⁸¹ See ICC Act s 5(3).

prosecution of the person before South African courts,⁸² although the Act is silent on what a court should consider in overruling such a decision. Further, while the ICC Act is also silent on the choice of court for prosecution, the Torture Act gives the NDPP the power to 'designate the court in which the prosecution must be conducted'.⁸³

D Jurisdiction in *SALC* and *al-Bashir*

Both the *SALC* and *al-Bashir* cases fell squarely within the subject-matter jurisdiction of the ICC Act, but raised different jurisdictional questions.

The jurisdictional issue in *al-Bashir* related to the reach of South Africa's enforcement jurisdiction. As the accused was present in South Africa after the commission of the alleged offences, it fits within s 4(3)(c) of the ICC Act. Moreover, the SCA in *al-Bashir* confirmed that national courts have jurisdiction to prosecute international crimes.⁸⁴ The case also raised the question of immunity, in the context of an arrest, as a procedural limitation to the exercise of enforcement jurisdiction, which we address below in Parts V and VI.

SALC raised the question of the exercise of universal jurisdiction in relation to the investigations of torture as a crime against humanity, a listed crime in the Rome Statute and the ICC Act.⁸⁵ Since the alleged crimes were committed outside of South African territory, in the absence of the 'traditional' jurisdictional connections, universal jurisdiction was the only possible basis to found the investigation. As confirmed by the Court, 'all states have an interest [in the crime of torture] under customary international law'.⁸⁶ However, for South Africa to exercise her jurisdiction, it must exclude the willingness and ability of Zimbabwe (where the crime occurred and which has primary jurisdiction on the basis of territoriality and nationality) to investigate the case.⁸⁷ Hence, as the *SALC* Court correctly observed, 'South African investigating institutions may investigate alleged crimes against humanity committed in another country by and against foreign nationals only if that country is unwilling or unable to do so itself'.⁸⁸ The Court noted that 'Zimbabwe was not asked by the alleged victims of torture to investigate the crime' but that 'it was unlikely that the Zimbabwean police would have pursued the investigation with the necessary zeal in view of the high profile personalities to be investigated'.⁸⁹ In fact, the lack of evidence pointing to Zimbabwe's courts launching an investigation indicated its unwillingness to do so.⁹⁰

⁸² See ICC Act s 5(6).

⁸³ Torture Act s 6(2).

⁸⁴ *Al-Bashir* (note 19 above) at para 1.

⁸⁵ See Rome Statute art 7; and ICC Act schedule 1.

⁸⁶ *SALC* (note 23 above) at para 49.

⁸⁷ This would be in line with Rome Statute arts 17(1)(a) and (b) (Require that, for a case to be admissible, the ICC must establish that the case is not being investigated or prosecuted by a state with jurisdiction over the case and the state is 'unwilling or unable genuinely to carry out the investigation or prosecution', or the case has been investigated by a state with jurisdiction over the case and the state has decided not to prosecute on grounds of inability and unwillingness).

⁸⁸ *SALC* (note 23 above) at para 62.

⁸⁹ *Ibid.*

⁹⁰ *Ibid* at para 78.

Further, the ICC Act limits prosecution based on universal jurisdiction to an accused who is present in South Africa. But is ‘presence’ required in the exercise of universal jurisdiction in the context of investigation of international crimes or is it limited to prosecution only? In *SALC*, the SAPS argued that ‘it has no duty to investigate the alleged torture in Zimbabwe because the suspects are not present in South Africa’. An investigation can only commence in terms of s 4(3) of the ICC Act, it argued, when an accused is in South Africa.⁹¹ The *SALC* Court held that the presence requirement is *not* applicable in relation to an investigation.⁹² This holding was based, first, on the right of an accused person ‘to be present when being tried’ in s 35(3)(e) of the Constitution which does not require ‘presence as a requirement for an investigation’.⁹³ Secondly, the Court considered s 4 of the ICC Act as well as international law scholarship and standards on the subject. It noted the lack of unanimity in international law scholarship on the issue of presence being a requirement for an investigation,⁹⁴ but referred, with approval, to paragraph 3(b) of the Resolution of the Institut de Droit International, which reads:

*Apart from acts of investigation and requests for extradition, the exercise of universal jurisdiction requires the presence of the alleged offender in the territory of the prosecuting State or on board a vessel flying its flag or an aircraft which is registered under its laws, or other lawful forms of control over the alleged offender.*⁹⁵

Hence, the exercise of universal jurisdiction by states includes investigative acts *in absentia* as well as requests for extradition.⁹⁶

The Court concluded that, while an accused has to be present before trial commences, ‘investigations in the absence of a suspect’ are permitted.⁹⁷ This approach is in line with the Rome Statute which draws a distinction between ‘investigation’ and ‘prosecution’ under art 17 and part V (arts 53–61) of the Statute. Put simply, the ‘exercise of universal jurisdiction, for purposes of the investigation of an international crime committed outside our territory, may occur in the absence of a suspect without offending our Constitution or international law’.⁹⁸ The Court justified its view as follows:

Requiring presence for an investigation would render nugatory the object of combating crimes against humanity. If a suspect were to enter and remain briefly in the territory of a state party, without a certain level of prior investigation, it would not be practicable to initiate charges and prosecution. ... Furthermore, any possible next step that could arise as a result of an investigation, such as a prosecution or an extradition request, requires an assessment of information which can only be attained through an investigation.⁹⁹

⁹¹ Ibid at paras 43–44.

⁹² Ibid at paras 43 and 47.

⁹³ Ibid.

⁹⁴ Ibid at para 46.

⁹⁵ Institut de Droit International *Universal Criminal Jurisdiction with regard to the Crime of Genocide, Crimes against Humanity and War Crimes* (2005), available at http://www.idi-iil.org/app/uploads/2017/06/2005_kra_03_en.pdf (emphasis in original quotation).

⁹⁶ See Kreß (note 44 above) at 576–578.

⁹⁷ See *SALC* (note 23 above) at para 46.

⁹⁸ Ibid at para 47.

⁹⁹ Ibid at para 48.

The Court emphasised that an investigation is important to the subsequent stages of prosecution or extradition. It held that 'South Africa *may*, through universal jurisdiction, assert prescriptive and, to some degree, adjudicative jurisdiction by investigating the allegations of torture as a precursor to taking a possible next step against the alleged perpetrators such as a prosecution or an extradition request'.¹⁰⁰

IV OBLIGATIONS IN RELATION TO THE INVESTIGATION AND THE PROSECUTION OF INTERNATIONAL CRIMES

South Africa has assumed various obligations in relation to the investigation and the prosecution of international crimes under international and domestic law. Its obligations can be placed into three categories: investigation, prosecution, and cooperation. While the obligation to investigate was at the core of *SALC*,¹⁰¹ the obligation to cooperate in the prosecution of international crimes was at issue in *al-Bashir*. The obligation to prosecute was not at issue in *SALC* and *al-Bashir* but the courts made reference to it.

In this Part, we discuss all three categories of obligations under both international and domestic law.

A Obligation to Investigate

1 International Law

The obligation to investigate international crimes stems from both customary and conventional international law. Under customary international law such an obligation can be derived from a combined reading of a number of sources (relevant treaty provisions, diplomatic practice, customary law on international crimes and the practice of tribunals under the rules of state responsibility).¹⁰² In relation to crimes against humanity, including torture, this obligation is accentuated by its customary law status.¹⁰³ More specifically, the ICTY in *Furundžija* noted that, in the context of criminal liability, the obligation on states to investigate acts of torture is a consequence of its customary international law and *jus cogens* status.¹⁰⁴ Similarly, the *SALC* Court held that 'the customary international law nature of the crime of torture underscores the duty to investigate this type of crime'.¹⁰⁵ As noted earlier, the *SALC* Court found that, in the context of the exercise of

¹⁰⁰ Ibid at para 49.

¹⁰¹ See *SALC* (note 23 above) at para 21.

¹⁰² See N Roht-Arriaza 'Nontreaty Sources of the Obligation to Investigate and Prosecute' in N Roht-Arriaza (ed) *Impunity and Human Rights in International Law and Practice* (1995) 40; and N Roht-Arriaza 'State Responsibility to Investigate and Prosecute Grave Human Rights Violations in International Law' (1990) 78 (2) *California Law Review* 449, 489.

¹⁰³ *Prosecutor v Tadić*, Case No. IT-94-1-T, ICTY Trial Chamber, Opinion and Judgment (7 May 1997) at paras 618–623 (Customary law status confirmed).

¹⁰⁴ See *Prosecutor v Furundžija*, Case No. IT-95-17/1-T, ICTY Trial Chamber Judgment (10 December 1998) at para 156.

¹⁰⁵ *SALC* (note 23 above) at para 60(b).

universal jurisdiction, the investigation of torture may proceed ‘in the absence of a suspect’, so long as international law is not breached.¹⁰⁶

In relation to conventional international law, the duty to investigate torture is, for example, grounded in art 5 of the CAT that require states to exercise jurisdiction over acts of torture. The ICTY has interpreted this provision to include ‘jurisdiction to *investigate*, prosecute and punish offenders’.¹⁰⁷ The duty on states parties in art 6(2) of the CAT to institute a preliminary inquiry further fortifies the existence of a legal duty under the CAT to investigate. Though art 5(2) of the CAT includes a presence requirement, Ventura has argued with reference to the *Lotus* case¹⁰⁸ that this does not prevent a state from proceeding with an investigation *in absentia* as long as it does not breach international law or domestic law.¹⁰⁹ In addition, in interpreting the relevant provision prohibiting torture in the International Covenant on Civil and Political Rights of 1966 (ICCPR), the UN Human Rights Committee has recognised the duty to investigate acts of torture as follows: ‘[c]omplaints must be *investigated promptly and impartially* by competent authorities’.¹¹⁰ Of course, not all investigations under the ICCPR would result in criminal prosecution.

While the Rome Statute does not contain a legal obligation on states parties to investigate international crimes, it recalls in its preamble ‘that it is the duty of every State to exercise its criminal jurisdiction over those responsible for international crimes’.¹¹¹ Also, if the ICC has concurrent jurisdiction over the offences, then the complementarity principle in the Rome Statute would imply that the primary obligation to investigate the listed international crimes rests on national jurisdictions.¹¹² In *SALC*, the Court confirmed, with reference to the complementarity principle in the Rome Statute, that ‘[t]he primary responsibility to investigate ... international crimes remains with states parties’.¹¹³ The ICC, however, in the absence of a UNSC referral, did not have jurisdiction in the *SALC* matter as it involved Zimbabwe, a non-state party to the Rome Statute. Despite this, the *SALC* Court emphasised the obligation on states parties to investigate international crimes, through the exercise of universal jurisdiction, committed within the territory, or by citizens, of non-state parties to the Statute.

¹⁰⁶ Ibid at para 47.

¹⁰⁷ *Furundžija* (note 104 above) at para 145 (emphasis added). See also Ventura (note 20 above) at 875–876 (Arguing with reference to art 5(2) of the CAT that ‘by necessary implication an *investigation* pursuant to universal jurisdiction is also an *obligation* under the treaty since an investigation must always occur before a prosecution is undertaken’).

¹⁰⁸ S.S. ‘*Lotus*’ (*France v Turkey*) (1927) PCIJ Series A, No 10 (‘*Lotus*’).

¹⁰⁹ Ventura (note 20 above) at 876.

¹¹⁰ Human Rights Committee *General Comment No. 20: Article 7 (Prohibition of torture, or other cruel, inhuman or degrading treatment or punishment)* (Forty-fourth Session, 1992), UN Doc. HRI/GEN/1/Rev.1 at 30 (1994) at paras 14–15.

¹¹¹ See MJ Ventura ‘The Duty to Investigate Zimbabwe Crimes against Humanity (Torture) Allegations: The Constitutional Court of South Africa speaks on Universal Jurisdiction’ (2015) 13 *Journal of International Criminal Justice* 861, 871–873 (Further discussion on understanding this preambular paragraph in the context of the duty to investigate in the exercise of universal jurisdiction).

¹¹² See Rome Statute arts 1, 17 and 18 read together.

¹¹³ *SALC* (note 23 above) at para 30.

This, the Court held, was necessary to prevent impunity, particularly where the non-state party has failed to institute an investigation.¹¹⁴

2 *Domestic Law*

The obligation to investigate international crimes also stems from domestic law. Under s 205(3) of the Constitution, the SAPS bears a constitutional obligation 'to prevent, combat and investigate crime'.¹¹⁵ And the ICC Act incorporates South Africa's obligations under the Rome Statute into domestic law. To meet these obligations, the SAPS Act establishes the Directorate for Priority Crime Investigation (the Hawks) as a unit under SAPS, with an obligation 'to prevent, combat and *investigate* national priority offences'.¹¹⁶ These are offences requiring 'national prevention or investigation' and include offences contained in Schedule 1 of the ICC Act, one of which is crimes against humanity.¹¹⁷

It is clear from *SALC* and the relevant domestic law, that the SAPS not only has a duty to investigate the crime of torture as a crime against humanity but also to prioritise it.¹¹⁸ The SAPS Act read with the NPA Act also envisage a cooperative role between the SAPS and the NPA in the investigation of crimes in the ICC Act.¹¹⁹

In *SALC*, the SAPS had advanced four main reasons for not proceeding with an investigation. Firstly, it lacked extraterritorial jurisdiction and the anticipated presence of the alleged perpetrators was not sufficient to trigger the required power and jurisdiction.¹²⁰ The Court found this to be a misconception of its duty as presence is not a requirement under international or domestic law.¹²¹ Secondly, a political justification – that an investigation would damage political relations between Zimbabwe and South Africa.¹²² The Court reasoned that such a justification undermines the principle of accountability for international crimes, especially as political tensions are often unavoidable.¹²³ Thirdly, it viewed the complainant as not impartial and that the complainant's assistance in the investigation could be seen as a 'covert agent' of SAPS, which would be at odds with the principle of state sovereignty.¹²⁴ The Court pointed out that SAPS's impartiality and not the complainant's is what matters.¹²⁵ Finally, investigation would be pointless because, as long as the accused are outside South Africa, South African courts do not have the jurisdiction to adjudicate.¹²⁶ Majiedt AJ avoided

¹¹⁴ Ibid at para 32.

¹¹⁵ Constitutional jurisprudence has also confirmed SAPS' duty to investigate crime. See the cases cited in *SALC* (note 23 above) at para 51 and fn 56 and 58.

¹¹⁶ SAPS Act ss 17C(1) and 17D(1)(a) (emphasis added).

¹¹⁷ Ibid ss 17A, 16(1) and 16(2)(iA), and item 4 of the Schedule.

¹¹⁸ See *SALC* (note 23 above) at para 57.

¹¹⁹ Ibid at para 58.

¹²⁰ Ibid at para 73.

¹²¹ Ibid.

¹²² Ibid at para 74.

¹²³ Ibid.

¹²⁴ Ibid at para 75.

¹²⁵ Ibid.

¹²⁶ Ibid at para 76.

this difficulty by holding that it fell outside the question in the case as it related to enforcement jurisdiction to prosecute and not investigations.¹²⁷

After rejecting these arguments against an obligation to investigate, the Court weighed several relevant considerations, including the principles of subsidiarity (whether there is ‘a substantial and true connection between the subject-matter and the source of the jurisdiction’)¹²⁸ and practicability (the ability to gather evidence and potentially prosecute) that are limitations to the duty to investigate international crimes. The Court was ultimately of the view that the SAPS did not act *reasonably* in not complying with its obligation to investigate the allegations of torture committed in Zimbabwe, as the threshold required to decline an investigation had not been met;¹²⁹ and that non-compliance was tantamount to tolerating impunity and providing a ‘safe haven’ for offenders.¹³⁰

B Obligation to Prosecute

1 *International Law*

Due to the customary law nature of core international crimes – such as crimes against humanity, genocide and war crimes – the obligation to prosecute constitutes an obligation that is owed to the international community as a whole.¹³¹ The international law principle of *aut dedere aut judicare* (extradite or prosecute) – which has its roots in customary international law and is reproduced in various treaties – reinforces the obligation of states to prosecute, particularly in situations where an accused is found in the territory of a state.¹³² This principle also promotes cooperation between states in the prosecution of international crimes.

In terms of treaty law, the obligation of states to prosecute torture is, for example, contained in the four Geneva Conventions, which place an obligation on states to prosecute grave breaches of the conventions.¹³³ Torture is listed as one of the acts that amount to a grave breach ‘if committed against persons or property protected by the Convention’.¹³⁴ The CAT, too, places an obligation on states parties to criminalise torture at the domestic level and ensure that acts of

¹²⁷ Ibid.

¹²⁸ Ibid at para 61.

¹²⁹ Ibid at paras 61–64 and 77–79.

¹³⁰ Ibid at para 80. See also *ibid* at paras 63–64 (Whether it is reasonable to proceed with an investigation, the Court held, should be considered on a case-by-case basis, taking into consideration relevant circumstances, such as: ‘whether the investigation is likely to lead to a prosecution and accordingly whether the alleged perpetrators are likely to be present in South Africa on their own or through an extradition request; the geographical proximity of South Africa to the place of the crime and the likelihood of the suspects being arrested for the purpose of prosecution; the prospects of gathering evidence which is needed to satisfy the elements of a crime; and the nature and the extent of the resources required for an effective investigation’).

¹³¹ Bantekas (note 1 above) at 379. See also R Memari ‘The Duty to Prosecute Crimes against Humanity under Universal Jurisdiction, Customary International Law, and Conventional International Law’ (2012) 31 *International Proceedings of Economics Development and Research* 131–132.

¹³² See Bantekas (note 1 above) at 378.

¹³³ Geneva Convention I art 49; Geneva Convention II art 50; Geneva Convention III art 129; Geneva Convention IV art 146; Additional Protocol I arts 11 and 85.

¹³⁴ Geneva Convention I art 50; Geneva Convention II art 51; Geneva Convention II art 130; Geneva Convention IV art 147.

torture are punished.¹³⁵ It also requires them, in the case of an alleged offender being within their jurisdiction, to 'submit the case to its competent authorities for the purpose of prosecution'.¹³⁶ Based on these and other provisions, the *SALC* Court¹³⁷ concluded that 'there is an international treaty law obligation to prosecute torture'.¹³⁸ A similar obligation exists with regard to genocide.¹³⁹

In relation to the Rome Statute, the complementarity principle again places the primary obligation to prosecute on national jurisdictions.¹⁴⁰ In fact, states parties have affirmed 'that the most serious crimes of concern to the international community as a whole must not go unpunished and that their effective prosecution must be ensured by taking measures at the national level and by enhancing international cooperation'.¹⁴¹

2 Domestic Law

National jurisdictions have a primary obligation to prosecute international crimes, as recognised in *SALC*.¹⁴² One of the objectives of the ICC Act is to enable the prosecution of persons that are accused of committing the recognised international crimes within South Africa, or in certain circumstances, outside South Africa.¹⁴³ The obligation to prosecute is however not mandatory, as the Act recognises that South African courts can decline to prosecute or be unable to prosecute.¹⁴⁴ A decision not to prosecute an accused 'does not preclude the prosecution of that person in the [ICC]' nor does it absolve South Africa from its obligation to cooperate with the ICC (discussed below).¹⁴⁵ The presence of an accused in the country in relation to crimes committed outside South Africa's territory is only a requirement for prosecution.¹⁴⁶ The Act further contains provisions on

¹³⁵ See CAT art 4.

¹³⁶ CAT art 7(1).

¹³⁷ Convention on the Prevention and Punishment of the Crime of Genocide (1948) (Genocide Convention).

¹³⁸ See *SALC* (note 23 above) at para 38, referring to CAT arts 4, 5 and 7, Geneva Conventions I to IV, and Genocide Convention arts 1, 2, 4 and 6.

¹³⁹ The Genocide Convention accentuates an obligation to prosecute the crime of genocide. Pursuant to arts 1 and 5 of the Convention, states parties have an obligation 'to prevent and to punish' genocide (whether committed in peacetime or wartime) and to 'provide effective penalties for persons guilty of genocide or any of the other acts enumerated in article III'. Under art IV, persons charged are to 'be tried by a competent tribunal of the State in the territory of which the act was committed, or by such international penal tribunal as may have jurisdiction'. Though the provision refers explicitly to territorial criminal jurisdiction, the application of universal jurisdiction to genocide is not prohibited as long as its application is consistent with (customary) international law. See *Application of the Convention on the Prevention and Punishment of Genocide (Bosnia Herzegovina v Serbia & Montenegro)* (2007) ICJ Reports at paras 442–443 ('*Bosnia Genocide*'); WA Schabas 'Introductory Note to the Convention on the Prevention and Punishment of the Crime of Genocide', available at <http://legal.un.org/avl/ha/cppcg/cppcg.html>; and ICRC 'Customary IHL Database - Rule 157. Jurisdiction over War Crimes', available at https://ihl-databases.icrc.org/customary-ihl/eng/docs/v1_cha_chapter44_rule157.

¹⁴⁰ See Rome Statute preamble and arts 1, 17–18, read together. See also *SALC* (note 23 above) at para 30.

¹⁴¹ See Rome Statute preamble.

¹⁴² See also *SALC* (note 23 above) at para 30.

¹⁴³ See ICC Act s 3(d).

¹⁴⁴ See ICC Act s 3(e).

¹⁴⁵ ICC Act ss 3(e) and 5(6).

¹⁴⁶ See ICC Act s 4(3)(c). See also *SALC* (note 23 above) at paras 41–49.

the procedure for the institution of prosecutions in South Africa, including the requirement of ‘consent’ from the NDPP, who in arriving at a decision must ‘give recognition to the obligation that the Republic, in the first instance and in line with the principle of complementarity as contemplated in Article 1 of the [Rome] Statute, has jurisdiction and the responsibility to prosecute persons accused of having committed a crime’.¹⁴⁷

Through the ICC Act, as well as the Constitution and the Torture Act, the South African government has complied with its obligation to criminalise torture.¹⁴⁸ An objective of the Torture Act is to provide for the prosecution of persons who commit torture.¹⁴⁹ The Torture Act, like the ICC Act, includes the ‘consent’ requirement of the NDPP for the prosecution of anyone who commits torture outside the territory of South Africa.¹⁵⁰

Lastly, the Geneva Conventions Act (GCA) includes an obligation to prosecute persons in relation to offences under the Act, one of which is torture (a grave breach of the Geneva Conventions). Section 7(1) of the GCA stipulates: ‘Any court in the Republic *may* try a person for any offence under this Act in the same manner as if the offence had been committed in the area of jurisdiction of that court, notwithstanding that the act or omission to which the charge relates was committed outside the Republic.’¹⁵¹ The obligation is, however, couched in discretionary language.

C Obligation to Cooperate

1 *International Law*

International tribunals are largely dependent on international cooperation in relation to investigation, arrest, prosecution and enforcement of their decisions.¹⁵² In international criminal law, state cooperation in the prosecution of international crimes can occur in two contexts: state-to-state cooperation (horizontal) or state-to-tribunal cooperation (vertical).¹⁵³ *Al-Bashir* related to the latter, hence our focus here in terms of obligations to cooperate is limited to vertical cooperation. In this context, the obligation to cooperate generally applies to the investigation and prosecution of international crimes, as well as to the post-prosecution phase (eg to cooperate in the implementation of a sentence through a state’s prison system, as international tribunals do not have their own permanent prisons). Furthermore, from an international law perspective, the source of the obligation to cooperate in the investigation and prosecution of international crimes before

¹⁴⁷ ICC Act s 5.

¹⁴⁸ See *S.A.L.C.* (note 23 above) at paras 38–39.

¹⁴⁹ See Torture Act s 2(1)(b).

¹⁵⁰ See Torture Act s 6(2).

¹⁵¹ Emphasis added.

¹⁵² See Cryer, Friman, Robinson & Wilmschurst (note 1 above) at 509.

¹⁵³ *Ibid.* It should be noted that the cooperation mechanism of the ICC is also seen by some as arguably horizontal on the basis that the ICC’s relationship with the international community is based on an agreement. See Bantekas (note 1 above) at 370.

the ICC could be the Rome Statute, the UN Charter, the Genocide Convention and/or an ad hoc agreement.¹⁵⁴

a Rome Statute

States parties to the Rome Statute and states that have accepted the ICC's jurisdiction pursuant to art 12(3) of the Statute have a general duty to 'cooperate *fully* with the Court in its investigation and prosecution of crimes within the jurisdiction of the Court'.¹⁵⁵ The obligation to cooperate includes the 'arrest and surrender' of persons against whom a warrant of arrest has been issued.¹⁵⁶ In the case of al-Bashir, two warrants of arrest were issued by the ICC. Following their issuance, the ICC confirmed that 'both warrants of arrest, together with cooperation requests for the arrest and surrender to the Court of Omar al-Bashir, have been transmitted ... to all States Parties to the Rome Statute, including the Republic of South Africa'.¹⁵⁷ On 28 May 2015, the South African government was reminded of its obligation to arrest and surrender al-Bashir and to consult the ICC in case of any difficulties in complying with the request.¹⁵⁸ States parties have to apply national procedures in enforcing a request for arrest and surrender¹⁵⁹ and therefore an obligation to 'ensure that there are procedures available under their national law for all of the forms of cooperation' under the Rome Statute.¹⁶⁰ Further, there are minimum requirements in the Rome Statute in relation to national arrest proceedings; and in the arrest process, states parties must act in accordance with both the Rome Statute and their domestic laws.¹⁶¹

The duty to cooperate is, however, couched in weak terms, as it is based on a 'request' (with the ICC having the powers to make requests of varying nature) as opposed to an 'order'. It is also subject to exceptions, as states parties have the option to 'deny' or 'postpone' the implementation of the request for cooperation on certain identified grounds. It can be 'denied' if it 'concerns the production of any documents or disclosure of evidence which relates to [a state's] national security'.¹⁶² The implementation of a cooperation request can be 'postponed' for as long as its 'immediate execution ... would interfere with an ongoing investigation or prosecution of a case different from that to which the request relates' or while the ICC is considering an admissibility challenge.¹⁶³ In addition, if the request 'is prohibited in the requested State on the basis of an existing

¹⁵⁴ As regards ad hoc cooperation agreements, see Rome Statute art 87(5) (Stipulates that, on the invitation of the ICC, non-states parties can sign such agreements with the ICC in relation to their provision of assistance to the ICC).

¹⁵⁵ Rome Statute art 86 (emphasis added).

¹⁵⁶ Rome Statute arts 58(5) and 89(1).

¹⁵⁷ *The Prosecutor v Omar Hassan Ahmad al-Bashir* (Decision Following the Prosecutor's Request for an Order Further Clarifying that the Republic of South Africa is under the Obligation to Immediately Arrest and Surrender Omar al-Bashir) ICC-02/05-01/09-242 (13 June 2015) PT Ch II at para 2.

¹⁵⁸ *Ibid* at para 3.

¹⁵⁹ See Rome Statute art 89(1).

¹⁶⁰ Rome Statute art 88.

¹⁶¹ See Rome Statute arts 58, 59(1) and 89(1).

¹⁶² Rome Statute art 93(4).

¹⁶³ Rome Statute arts 94–95.

fundamental legal principle of general application¹⁶⁴ and the matter could not be resolved in consultations, the ICC ‘*shall modify*’ the request as necessary.¹⁶⁵ The above exceptions, as well as the issue of competing requests for surrender,¹⁶⁶ were not at issue in *Bashir*.

Of relevance for our present purposes is the exact meaning of art 98 of the Rome Statute, which stipulates:

1. The Court *may not proceed* with a request for surrender or assistance which would require the requested to act inconsistently with its obligations under international law with respect to the State or diplomatic immunity of a person or property of a third State, unless the Court can first obtain the cooperation of that third State for the waiver of the immunity.
2. The Court *may not proceed* with a request for surrender which would require the requested State to act inconsistently with its obligations under international agreements pursuant to which the consent of a sending State is required to surrender a person of that State to the Court.¹⁶⁷

Article 98(1) relates to immunities (diplomatic or other) under customary or conventional international law and applies in the context of a request for ‘surrender’ or ‘assistance’, whereas art 98(2) concerns extradition treaties or exclusive jurisdiction agreements and applies only in the context of a request for ‘surrender’.¹⁶⁸ In both instances, the requested state is in a conflicting situation.¹⁶⁹ Such conflicts should and could be avoided if the ICC adequately addresses them when issuing a warrant of arrest. This was not done when the warrant of arrest against al-Bashir was issued, resulting in issues of immunity and conflicting obligations subsequently being raised by several states parties, including Kenya and South Africa.

It should first be noted that both subparagraphs are explicitly directed to the ICC ([t]he Court may not proceed) and not – unlike, for example, arts 93(4) and 95 of the Rome Statute (‘a State party may deny’ and ‘the requested State may postpone’) – to states parties. Therefore, art 98 of the Rome Statute does not include a right for the requested state to refuse to execute a request for arrest and surrender once it is made.¹⁷⁰

Secondly, some scholars assert that art 98 of the Rome Statute is ‘formulated in such a way as to limit the power of the Court in the matter of request for surrender and assistance’,¹⁷¹ and argue for an obligation ‘not to put a state in

¹⁶⁴ Rome Statute art 93(4).

¹⁶⁵ Ibid (emphasis added).

¹⁶⁶ See Rome Statute art 90. Note that states parties are required to ‘promptly consult’ with the ICC.

¹⁶⁷ Emphasis added. It should be noted that because the Rome Statute is silent on whether the competing international agreement referred to in art 98(2) should precede the Rome Statute, states have gone further to adopt bilateral impunity agreements. It has, however, been argued that post-impunity agreements entered into by states parties to the Statute would amount to a breach of their obligations under the Statute. See Bantekas (note 1 above) at 439.

¹⁶⁸ See Bantekas (note 1 above) at 439; and Cryer, Friman, Robinson & Wilmschurst (note 1 above) at 512.

¹⁶⁹ This conflict is explored further in the subsequent Sub-Part IV.C.3 and Part VI of this article.

¹⁷⁰ Member states are, however, permitted, as Bantekas argues, to depart from the obligation to assist or surrender to the court, in situations where a multiple, competing request is premised on a treaty or customary obligation with a third party. Bantekas (note 1 above) at 439.

¹⁷¹ P Gaeta ‘Does President al-Bashir Enjoy Immunity from Arrest?’ (2009) 7(2) *Journal of International Criminal Justice* 328.

the position of having to violate its international obligations with respect to immunities'.¹⁷² Whilst the phrase 'shall' constitutes an obligation to do something and 'shall not' an obligation to not do something – both denoting 'an imperative command' – 'may' means either '[t]o be permitted to', or '[t]o be a possibility' or '[l]oosely, is required to; shall; must'.¹⁷³ Looking at the relevant case law, 'it is only by considering the general provisions of the law in question and the purview of the whole legislation on the subject that we can tell whether "may" confers a discretionary power or imposes an obligatory duty'.¹⁷⁴ Put differently, in statutes, the word 'may' must be read in context to determine if it means an act is optional or mandatory, for it may be an imperative. For our purposes, the problematic question is the exact meaning of the phrase 'may not (proceed)' in the Rome Statute. On the one hand, it could indeed indicate that there is *no permission* (or *no possibility*) to proceed, and on the other hand, it could refer to a *permission* (or *possibility*) *in terms of a choice* to either proceed or not proceed.

The first approach – *no permission* to proceed – is in line with the relevant provision in the French and Spanish versions of the Rome Statute, which if translated literally, both mean 'not being allowed to'.¹⁷⁵ Moreover, since the ICC is not obligated to request assistance but rather has 'the authority to make requests',¹⁷⁶ i.e. the permission to do so, the first approach seems legally meaningful. The implication of such an interpretation – *no permission* to proceed – would be that, if the ICC has not obtained a waiver of immunities from the third state, the requested state would not commit an international wrongful act if it refuses to cooperate with the ICC. In the case of al-Bashir, the ICC has not obtained a waiver of immunities from Sudan¹⁷⁷ and would therefore not be allowed to proceed with the request for cooperation. Thus, from an international law perspective, the South African government could have lawfully disregarded the request.¹⁷⁸ Such an understanding would be based on an assumption that the granting of a waiver must be obtained from the third state, in this case Sudan,

¹⁷² K Prost & A Schlunck, 'Article 98' in O Triffterer (ed) *Commentary on the Rome Statute of the International Criminal Court: Observers' Notes. Article by Article* (1999) 1131. See also Gaeta (note 171 above) at 315 (Concluding that states parties are not bound to comply with a request if ultra vires with art 98); and JD van der Vyver 'Al-Bashir Debacle' (2015) 15 *African Human Rights Law Journal* 570 ('precludes the ICC from proceeding').

¹⁷³ *Black Law Dictionary* (8th Edition, 2004) 1000. For case law confirming the permissive nature of 'may' and on 'may' being synonymous with 'shall' or 'must' in an effort to effectuate legislative intent, see *South African Legal Dictionary* (3rd Edition, 1951) 482; and JB Saunders *Words and Phrases Legally Defined* (3rd Edition, 1988–1990) 342 et seq.

¹⁷⁴ Saunders (note 173 above) at 482.

¹⁷⁵ Rome Statute art 128 stipulates: 'The original of this Statute, of which the Arabic, Chinese, English, French, Russian and Spanish texts are equally authentic ...'. The French text reads: 'La Cour ne peut poursuivre l'exécution d'une demande de remise ou d'assistance'; the Spanish text reads: 'La Corte no dará curso a una solicitud de entrega o de asistencia'; and the German text reads 'Der Gerichtshof darf kein Überstellungs- oder Rechtshilfeersuchen stellen'.

¹⁷⁶ Rome Statute art 87.

¹⁷⁷ See the Part IV.A on immunities below (Only Sudan can grant such a waiver).

¹⁷⁸ See also Gaeta (note 171 above) (Concludes that the request to surrender is not issued in accordance with the Rome Statute and thus not binding on states parties).

and can neither expressly nor implicitly be waived by another actor, such as by the UNSC through a referral.¹⁷⁹

One could, however, contend that in conformity with the interpretation of the phrase ‘shall’ (*obligation to do something*) and ‘shall not’ (*obligation to not do something*), the *permission* or *possibility in terms of a choice* (to do or not do something) is a key element for interpreting ‘may’ or ‘may not’, including the effect of leaving the finalisation of such a decision at the judicial discretion of the ICC. This discretion is fortified by art 119(1) of the Rome Statute, requiring that ‘[a]ny dispute concerning the judicial functions of the Court shall be settled by the direction of the Court’. Let us assume that the drafters of the Rome Statute were aware that the ICC is not obligated to request assistance but has ‘the *authority* to make requests to States Parties for cooperation’.¹⁸⁰ And assume that they did not discount this nor want to include legally ineffective wording. Therefore, the fact that art 98 of the Rome Statute reiterates this ‘authority’ could indicate an obligation on the part of the ICC to seriously consider the conflicting obligations in relation to third countries before exercising its discretion whether to proceed or not to proceed.¹⁸¹ The implication of this understanding would be that, if the ICC insists on the request after considering the conflict, conflicting obligations cannot be grounds for a refusal to execute a request for arrest and surrender.¹⁸² Thus, if a requested state persists in asserting that a conflict exists amidst the ICC’s insistence, the result could be non-compliance proceedings.¹⁸³ In relation to al-Bashir, the ICC stated that South Africa had no conflicting obligations due to the implicit waiver of al-Bashir’s immunity by UNSC Resolution 1593 (2005), and insisted on its request.¹⁸⁴ Therefore, irrespective of an express waiver of immunities by Sudan, the South African government would have been obligated to execute the ICC’s request.

Another consideration is that the drafters of the Rome Statute identified, in art 97, difficulties that would hamper or prevent a state party’s implementation of a cooperation request. These are: (a) the information is not sufficient to execute the request; (b) the person against whom a warrant has been issued cannot be located, the person named in the arrest warrant is not the same person that is in the requested state; and (c) its execution ‘would require the requested State to breach a pre-existing treaty obligation undertaken with respect to another State’. The introductory phrase ‘*inter alia*’ clarifies that this is not an exhaustive list. A state faced with difficulties must consult with the ICC without delay so that the challenge can be resolved.¹⁸⁵ Hence, the South African government consulted

¹⁷⁹ The question of whether the UNSC is permitted to waive immunities, and whether it has, through its referral, waived al-Bashir’s immunity is considered in Part IV.A below.

¹⁸⁰ Rome Statute art 87(1)(a)(emphasis added).

¹⁸¹ This could, for example, include procedural obligations to document its considerations.

¹⁸² See Cryer, Friman, Robinson & Wilmshurst (note 1 above) at 513.

¹⁸³ Ibid.

¹⁸⁴ ICC-02/05-01/09-242 (note 157 above) at paras 4–7 (Stating that SA had no conflict/competing obligations due to implicit waiver of immunity by UNSC Resolution 1593), with reference to *Prosecutor v Omar Hassan Ahmed al-Bashir (Decision on the Cooperation of the Democratic Republic of the Congo Regarding Omar al-Bashir’s Arrest and Surrender to the Court)* ICC-02/05-01/09-195 (9 April 2014) PT Ch II, para 28–31.

¹⁸⁵ See art 97 of the Rome Statute.

with the ICC indicating that it 'was subject to competing obligations' and 'there was a lack of clarity in the law'.¹⁸⁶ To which the ICC indicated that 'there was no ambiguity in the law' and that the South African government had an obligation to immediately arrest and surrender al-Bashir to the ICC as soon as he was, at that time, on South African territory.¹⁸⁷ The ICC added that the consultations did 'not trigger any suspension or stay' of South Africa's obligation to cooperate.¹⁸⁸

It is worth noting that the government's approach from 2015 – invoking conflicting obligations and a lack of clarity in the law – contradicts its approach in 2009 on the same matter, where it confirmed its obligation to arrest al-Bashir. For example, in its submission to the ICC on the question of non-compliance with its obligation to cooperate in the arrest and surrender of al-Bashir, South Africa cited the 'dilemma' it was placed in relation to the 'peace-justice relationship'. It also argued that it did not fail to comply with its obligations due to the immunity that Al-Bashir enjoyed which had not been expressly waived by Sudan or implicitly waived by the UNSC through its referral resolution, thus precluding the request for cooperation by virtue of article 98 of the Rome Statute.¹⁸⁹ In contrast, in 2009, the country's officials, following the invitation of al-Bashir to attend the South African president's inauguration, confirmed that he would be arrested upon his arrival in the country, in execution of the ICC's warrants of arrest.¹⁹⁰ This resulted in al-Bashir declining the invitation.¹⁹¹

b UN Charter

As noted previously, obligations to cooperate in the investigation or prosecution of international crimes under the ICC regime can also stem from the UN Charter through a UNSC referral resolution. The UNSC is empowered, acting under Chapter VII of the UN Charter, to refer a situation to the ICC where the crimes in the Statute appear to have been committed.¹⁹² While acting under Chapter VII, the referral can be made in relation to any UN member state even if the state is not a state party to the Rome Statute. This was the case with Sudan: the situation in Darfur was referred to the ICC by the UNSC.¹⁹³

In making such a referral, the UNSC can impose cooperation obligations that are binding by virtue of art 25 (read with art 24(1)) of the UN Charter.¹⁹⁴ In the case of Sudan, the UNSC decided that 'the Government of Sudan and all other parties to the conflict in Darfur, *shall cooperate fully with and provide any necessary*

¹⁸⁶ ICC-02/05-01/09-242 (note 157 above) at para 4.

¹⁸⁷ Ibid at paras 5, 8, 9 and 10.

¹⁸⁸ Ibid at para 8.

¹⁸⁹ *Prosecutor v Omar Hassan Ahmad Al Bashir* (Submission from the Government of the Republic of South Africa for the purposes of proceedings under Article 87(7) of the Rome Statute) ICC-02/05-01/09-290 (17 March 2017) PT Ch II, paras 20 and 52.

¹⁹⁰ See *al-Bashir HC* (note 24 above) at para 12.

¹⁹¹ Ibid.

¹⁹² See Rome Statute art 13(b).

¹⁹³ UNSC Resolution 1593 (note 70 above) at para 1.

¹⁹⁴ UN Charter art 25 stipulates: 'The Members of the United Nations agree to accept and carry out the decisions of the Security Council in accordance with the present Charter.' Article 24(1) gives the UNSC primary responsibility for the maintenance of international peace and security and for it to act on behalf of UN member states.

assistance to the Court and the Prosecutor'.¹⁹⁵ Irrespective of being a party to the Statute, it further 'urges all States and concerned regional and other international organizations to cooperate fully', while 'recognising that States not party to the Rome Statute have no obligation under the Statute'.¹⁹⁶ The resolution is, evidently, worded narrowly as the obligation to cooperate fully with the ICC is directed to Sudan¹⁹⁷ and other parties to the conflict, and not to all UN member states. States not party to the conflict, such as South Africa, and international organisations are merely *urged* to cooperate fully. As Akande rightly explains, '[a]n urging to cooperate is manifestly not intended to create an obligation to do so' and '[t]he word "urges" suggests nothing more than a recommendation or exhortation to take certain action'.¹⁹⁸

c Genocide Convention

In relation to the charge of genocide in al-Bashir's case, it has been argued that an obligation to cooperate also stems from the Genocide Convention.¹⁹⁹ Article VI of the Convention allows for the trial of persons charged with genocide by an international penal tribunal. In *Bosnia Genocide*, the ICJ held that states parties to the Convention have an obligation to cooperate with the international penal tribunal and that art VI of the Convention

obliges the Contracting Parties 'which shall have accepted its jurisdiction' to co-operate with it, *which implies that they will arrest persons accused of genocide who are in their territory* – even if the crime of which they are accused was committed outside it – and, failing prosecution of them in the parties' own courts, that they will hand them over for trial by the competent international tribunal.²⁰⁰

Pursuant to the ICJ's reasoning, in our present case, it is important to first establish if the ICC qualifies as an 'international penal tribunal' within the meaning of art VI of the Convention.²⁰¹ According to the ICJ, such tribunal includes all international criminal courts created after the Convention was adopted and 'of potentially universal scope, and competent to try the perpetrators of genocide

¹⁹⁵ UNSC Resolution 1593 (note 70 above) at para 2 (emphasis added).

¹⁹⁶ Ibid (emphasis added).

¹⁹⁷ See *Prosecutor v Omar Hassan Ahmad al-Bashir* (Decision on the Prosecution's Application for a Warrant of Arrest against Omar Hassan Ahmad al-Bashir), ICC-02/05-01/09-3 (4 March 2009) PT Ch I at paras 240–249 (Referring to Sudan's obligation to fully cooperate with the ICC pursuant to resolution 1593).

¹⁹⁸ CD Akande 'The Legal Nature of Security Council Referrals to the ICC and its Impact on al-Bashir's Immunities' (2009) 7 *Journal of International Criminal Justice* 333, 344–345 (Further explains that such a recommendation does not fall within the category of recommendations that would come within the scope of art 103, as it is not an authorisation to states to take action under chapter VII of the UN Charter).

¹⁹⁹ See generally M Gillett 'The Call of Justice: Obligations under the Genocide Convention to Cooperate with the International Criminal Court' (2012) 23(1) *Criminal Law Forum* 63–96 ('[C]oncludes that there is an inherent obligation in the Genocide Convention to cooperate with international proceedings on genocide charges and that this obligation has been activated by Resolution 1593'). See also D Jacobs 'The Frog that Wanted to Be an Ox: The ICC's Approach to Immunities and Cooperation' in C Stahn (ed) *The Law and Practice of the International Criminal Court* (1st Edition, 2015) 296–297; and Akande (note 198 above) at 348–351.

²⁰⁰ *Bosnia Genocide* (note 139 above) at para 443.

²⁰¹ Ibid at para 444.

or any of the other acts enumerated in Article III.²⁰² The ICC meets this requirement and has jurisdiction over the crime of genocide as defined in the Genocide Convention.²⁰³

The second question is whether the state party concerned can be regarded as having 'accepted the jurisdiction' of the ICC within the meaning of art VI, which according to the ICJs reasoning, must consequently be formulated as whether the state party is 'obliged to accept the jurisdiction of the [ICC], and to co-operate with the Tribunal by virtue of the Security Council resolution ... or some other rule of international law'.²⁰⁴ Thus, South Africa, as a state party to the Genocide Convention that has accepted the jurisdiction of the ICC, has an obligation to cooperate under the Convention in the arrest of al-Bashir when he is on South African territory. Sudan, as a state party to the Genocide Convention, that is obliged to accept the jurisdiction of the ICC by virtue of resolution 1593, has an obligation under the Convention to cooperate with the ICC. For other non-state parties to the Rome Statute, who are parties to the Genocide Convention, considering that resolution 1593 merely *urges* them to cooperate, an obligation to cooperate with the ICC cannot be derived from the Convention unless it is established that 'some other rule of international law' obliges them to cooperate. This obligation to cooperate on South Africa and Sudan under the Convention applies irrespective of a person's official capacity since art VI of the Convention requires that even 'constitutionally responsible rulers, public officials or private individuals' who commit genocide or acts listed in art II must be punished.²⁰⁵

2 *Domestic Law*

South Africa's cooperation obligation can also stem from domestic law, more precisely the ICC Act. One of the objectives of the ICC Act is to provide for the government's cooperation with the ICC, particularly, amongst other things, in enabling the ICC to make assistance requests and providing mechanisms for the surrender, to the ICC, of persons accused of committing crimes under the Rome Statute.²⁰⁶ The obligation to cooperate in the ICC Act also includes arrests, and the Act provides a cooperation mechanism on the arrest and surrender of persons to the ICC, including procedures to be followed upon receipt of a warrant of arrest.²⁰⁷ Leaving aside for now the issue of immunities, South African authorities are thus required, in terms of the Act, to cooperate with the ICC in effecting the arrest of persons suspected of crimes under the Act.

²⁰² Ibid at para 445.

²⁰³ See also Jacobs (note 199 above) at 297–298 (Further reading, not disputing that the ICC can be seen as an international penal tribunal but also critiquing its application in relation to the ICC).

²⁰⁴ *Bosnia Genocide* (note 139 above) at paras 444 and 446 (The court referred specifically to UNSC resolution that establishes the tribunal but in applying the reasoning to our present case, one has to consider UNSC referral resolution instead which triggered ICC's jurisdiction over the Darfur situation).

²⁰⁵ On immunity under the Genocide Convention, see Part V.A.1.

²⁰⁶ See ICC Act preamble and s 3(e).

²⁰⁷ See ICC Act ss 10–13. Chapter 4 of the ICC Act generally provides a mechanism for the South African government's cooperation with the ICC.

Overall, provided that one accepts that ‘may not proceed’ leaves the ICC with the discretion on whether to request assistance or not, South Africa has an obligation to cooperate in the arrest and surrender of al-Bashir under both international and domestic law,²⁰⁸ according to the ICC Pre-Trial Chamber, without any basis under the Rome Statute for its postponement or refusal.

On 14 and 15 June 2015, al-Bashir was on South African territory for the AU Assembly’s 25th ordinary session,²⁰⁹ thus triggering South Africa’s cooperation obligations in relation to his arrest and surrender since the ICC had issued two warrants of arrest against him, including a formal request from the ICC for his arrest and surrender.²¹⁰ However, the government did not arrest him, despite a High Court order to this effect, because it contended that al-Bashir enjoys immunity from arrest.²¹¹ Hence, it was not disputed at the SCA that the government had an obligation to cooperate with the ICC but the government was of the view that the obligation is limited by the issue of immunity.²¹² The immunity question is discussed in parts V and VI of this article.

3 *AU Decision to not Cooperate*

After the indictment of President al-Bashir, the AU, at a meeting held in July 2009, endorsed a decision of African state parties to the Rome Statute proclaiming that ‘the AU member states shall not cooperate pursuant to the provisions of Article 98 of the Rome Statute of the ICC relating to immunities, for the arrest and surrender of President Omar El Bashir of The Sudan’.²¹³ Despite the lack of a provision in the Constitutive Act of the AU on the binding nature of AU decisions, failure to comply would attract sanctions.²¹⁴ In order to avoid sanctions, South Africa, as an AU member state, must comply with this decision. But that decision is in direct conflict²¹⁵ with South Africa’s international and domestic law obligation to cooperate in al-Bashir’s arrest and surrender.

a Conflicts between the obligation towards the AU and other international obligations to cooperate

Article 103 of the UN Charter serves to solve issues of conflicting treaty obligations incumbent upon UN members in favour of those stemming from the UN Charter. Cooperation obligations deriving from the UNSC referral would therefore prevail over obligations deriving from the AU Constitutive Act.

²⁰⁸ See *al-Bashir* (note 19 above) at paras 57, 58, 61, 65 and 113 (South Africa’s international law obligations) and at paras 86–105 and 113 (South Africa’s domestic law obligations).

²⁰⁹ Ibid at para 2.

²¹⁰ See note 21 above.

²¹¹ See *al-Bashir* (note 19 above) at para 4.

²¹² Ibid at para 65.

²¹³ See note 14 above.

²¹⁴ Article 23(2) of the Constitutive Act of the African Union (2000), OAU Doc CAB/LEG/23.15 (Stipulates that ‘any Member State that fails to comply with the decisions and policies of the Union may be subjected to other sanctions, such as the denial of transport and communications links with other Member States, and other measures of a political and economic nature to be determined by the Assembly’). See also Du Plessis & Gevers (note 15 above) at 1 (Arguing that, based on art 23 and the doctrine of implied powers, AU Assembly decisions ‘are potentially binding on member states’).

²¹⁵ On the notion of a norm conflict, see note 299 below.

In *al-Bashir*, however, South Africa's international cooperation obligations stem from the Rome Statute and/or the Genocide Convention, not from the UNSC referral. The referral merely *urged* the South African government to cooperate fully.²¹⁶ Therefore, conflicts between the Rome Statute and Genocide Convention obligations to cooperate on the one hand, and obligations towards the AU to not cooperate on the other hand, cannot be solved by art 103 of the UN Charter. If the UNSC had not *urged* South Africa to cooperate but *obliged* it to do so, then art 103 would have been applicable in solving the conflict.

The relations between rules generated by the two different treaties are governed by the three general principles on conflict resolution 'which in all legal orders regulate the relations between norms deriving from the *same source*':²¹⁷ a later law repeals an earlier one (*lex posterior derogat legi priori*); a later law, general in character, does not derogate from an earlier one, which is special in character (*lex posterior generalis non derogat priori speciali*); a special law prevails over a general law (*lex specialis derogat legi generali*).²¹⁸

Neither the obligation to cooperate in relation to the arrest and surrender of al-Bashir nor the obligation to not cooperate with the ICC in this regard is more general or special in character vis-à-vis the other.²¹⁹ The generality and speciality of a rule is always relational to some other rule since 'every general rule is particular, too, in the sense that it deals with some particular substance, that is, includes a certain fact-description as a *general* condition of its application'.²²⁰ In *al-Bashir*, none of the obligations in question are general in character. They are rather specific and include directly opposite instructions regarding the same subject matter, the arrest and surrender of al-Bashir, demanding of the South African government on the one hand to arrest and surrender al-Bashir to the ICC and, on the other hand, to set aside this obligation and not cooperate.

Moreover, South Africa's obligation towards the AU existed since July 2009, following the AU Assembly decision. The first formal cooperation request to all states parties by the Pre-Trial Chamber I for the arrest and surrender of al-Bashir to the ICC regarding crimes against humanity and war crimes, triggering the cooperation obligations of South Africa under the Rome Statute, was issued in March 2009.²²¹ Hence, in relation to the above two crimes, its obligation towards the AU would be the later one. In contrast, in relation to three counts of genocide, the obligation triggered by the second formal request was issued on 21 July

²¹⁶ See part IV.C.1.b above. But see Gaeta (note 171 above) at 326 et seq (Argues that the UNSC referral binds all UN member states to fully cooperate).

²¹⁷ A Cassese *International Law* (2nd Edition, 2005) ch 8, 154.

²¹⁸ For a detailed discussion of the various approaches on the legal nature of these rules, see, eg, E Vranes 'Lex Superior, Lex Specialis, Lex Posterior – Zur Rechtsnatur der "Konfliktlösungsregeln"' (2005) 65 *Heidelberg Journal of International Law* 391. On the legal nature of *lex specialis*, see, eg, S Zorzetto 'The Lex Specialis Principle and its Use in Legal Argumentation. An Analytical Inquire' (Sept 2012–Feb 2013) 3 *Ennomia. Revista en Cultura de la Legalidad* 61.

²¹⁹ We of course acknowledge that if the question of speciality relates to the treaty and its subject matter as a whole, as opposed to a particular rule, one could argue that, in contrast to the AU obligation, the obligation in the Rome Statute is more specific because it flows from a treaty dealing specifically with international crimes and its prosecution.

²²⁰ M Koskeniemi *Fragmentation of International Law: Difficulties Arising from the Diversification and Expansion of International Law* ILC Report of the Study Group A/CN.4/L.682 (13 April 2006) 111.

²²¹ See ICC-02/05-01/09-7 (note 21 above).

2010²²² and would therefore be the later law. Thus, applying the *lex posterior* principle, while in relation to crimes against humanity and war crimes South Africa's obligation towards the AU to not cooperate with the ICC would prevail, with respect to three counts of genocide these obligations would be subordinated.

An absurdity potentially occurs when either no precise date can be assigned to the creation of an obligation, for example, due to their gradual development (eg customary law or general principles)²²³ or the precise date is dependent on a further requirement that triggers the creation of the obligation (eg South Africa's cooperation obligation with the ICC in relation to al-Bashir). Such an absurdity also illustrates the limits of the *lex posterior* rule to resolve conflicts. In *Bashir*, the South African government's defiance could result in an even greater absurdity, if, for example, the ICC Pre-Trial Chamber I simply issues a new formal request in relation to crimes against humanity and war crimes, which then would be the later law and prevail. Thus, it is problematic to apply the *lex posterior* rule to solve the conflict between the obligation towards the AU to not cooperate with the ICC and the obligation to cooperate stemming from the Rome Statute (and from the Genocide Convention).

This line of reasoning is consistent with art 30(3) of the Vienna Convention on the Law of Treaties (VCLT), which 'effectively codifies the *lex posterior* rule',²²⁴ requiring that this rule only applies in situations where there are either identical parties in the later treaty or, in addition to all the parties of the earlier treaty, new state parties. Not all AU members are parties to the Rome Statute, nor vice versa. The same applies to the Genocide Convention. In the absence of any other international law rule governing the relation between international and regional obligations, both obligations remain equal in ranking.

b Conflicts between the obligation towards the AU and domestic obligations to cooperate

The conflicting international (at a regional (AU) level) and domestic law principles in relation to South Africa's obligation to cooperate with the ICC operate on different levels. Domestic laws cannot be invoked as justification for a state's failure to comply with its international obligations.²²⁵ International obligations, in turn, are only relevant before domestic courts to the extent provided for by domestic law.

²²² See note 21 above.

²²³ See M Akehurst 'Hierarchy of Sources' (1974/75) 47 *British Yearbook for International Law* 273 (Saying that in this situation, it is 'difficult to apply' the *lex posterior* rule). On the application of the *lex posterior* rule between treaties and customs see also Part VI in this article.

²²⁴ CJ Borgen 'Resolving Treaty Conflicts' (2005) 37 *George Washington International Law Review* 573, 603; and JB Mus 'Conflicts between Treaties in International Law' (1998) XLV *Netherlands International Law Review* 219, 220.

²²⁵ While art 27 of the VCLT applies to treaty obligations, art 32 of the Draft Articles of State Responsibility applies to customary international law obligations. This is supported by state practice and international decisions. See, in detail *Draft Articles on Responsibility of States for Internationally Wrongful Acts, with commentaries* (2001) 94, available at http://legal.un.org/ilc/texts/instruments/english/commentaries/9_6_2001.pdf). But see *al-Bashir* (note 19 above) at para 103 (SCA assumes that SA is 'entitled to depart from [customary international law] by statute as stated in s 232 of the Constitution').

V IMMUNITIES OF SENIOR STATE OFFICIALS AND INTERNATIONAL CRIMES

A International Law

It is an established principle under international law that incumbent heads of state or government and senior government officials enjoy personal immunity from criminal jurisdiction.²²⁶ While this does not include immunity from investigation and is thus not applicable in *SALC*, it comprises immunity from arrest and from the exercise of criminal jurisdiction.²²⁷ What this means is that domestic authorities cannot entertain a particular criminal suit; 'not that the defendant is discharged from criminal liability altogether or that the jurisdiction of the court is extinguished'.²²⁸ It is thus a 'procedural bar' and once removed, 'criminal liability of the accused re-emerges and that person becomes once again susceptible to criminal prosecution'.²²⁹

While state or sovereign immunity, in the context of criminal law, has been given a broad meaning to include head of state immunity,²³⁰ immunities have progressed in international law from absolute to restrictive. In addition, the doctrine of sovereign immunity has weakened considerably, particularly with respect to acts that constitute international crimes.²³¹ Further, while there has been a shift in priorities in favour of non-impunity, accountability and justice, and arguments have been advanced against immunity on the basis of the *jus cogens* nature of the prohibition of international crimes, international criminal law has not totally displaced international law relating to immunities, especially not in relation to national criminal proceedings in foreign states.²³² The position in international criminal law proceedings is thus different from that in national criminal proceedings.

1 *International Law in international criminal proceedings at the ICC*

The Rome Statute establishes a two-tier immunity structure for the ICC – one for officials from states parties and the other for officials from non-states parties.

²²⁶ See *Arrest Warrant* (note 50 above) at para 51; recently confirmed in the *Jurisdictional Immunities of the State (Germany v Italy)* (2012) ICJ Reports 99 at para 58 (*Jurisdictional Immunities*).

²²⁷ See in detail Gaeta (note 171 above) at 320.

²²⁸ Bantekas (note 1 above) 127, referring to *Dickinson v Del Solar* (1930) 1 KB 376, 380, per Lord Hewart CJ and *Arrest Warrant* (note 50 above) at paras 47–55.

²²⁹ Bantekas (note 1 above) at 127.

²³⁰ See D Tladi 'The Duty on South Africa to Arrest and Surrender President Al-Bashir under South African and International Law: A Perspective from International Law' (2015) 13(5) *Journal of International Criminal Justice* 1027, 1044 (Cites C Kreß, who argues for a broad interpretation of state immunity on the basis that a 'state' cannot be arrested and surrendered. Whether a narrow or broad interpretation of state immunity is applied under the Rome Statute is important, as art 98(1) refers to 'State or diplomatic immunity of a person or property'. As Tladi (at 1043–1044) explains, a narrow interpretation would imply that the exception in art 98(1) does not apply to al-Bashir as he is neither a state nor a diplomat (assuming also that he was not granted diplomatic status during his visit to South Africa) while a broad interpretation would allow for application of the exception).

²³¹ For a detailed discussion on this, see Bantekas (note 1 above) at 128 et seq. See also, generally, R van Alebeek *Immunities of States and their Officials in International Criminal Law and International Rights Law* (2008); and Y Naqvi *Impediments to Exercising Jurisdiction over International Crimes* (2010).

²³² See Cryer, Friman, Robinson & Wilmschurst (note 1 above) at 531–532.

On the one hand, in terms of art 27 of the Rome Statute, states parties accept that immunities do not bar ICC prosecution.²³³ Put differently, states parties cannot raise immunity of their former or their incumbent heads of states in proceedings before the ICC.²³⁴ On the other hand, in terms of art 98(1) of the Rome Statute, as stated above, the ICC ‘may not’ proceed with a request for surrender if it requires ‘the requested State to act inconsistently with its obligations under international law with respect to the State or diplomatic immunity of a person or property of a third State’, unless the third State waives the immunity. The two tiers are not contradictory since the first tier governs the ICC’s exercise of jurisdiction over an accused person before it, while the second tier applies only in the context of states parties’ obligations to cooperate with the ICC in the context of a request for surrender of incumbent heads of non-states parties.²³⁵

Generally, the doctrine of sovereign immunity derives from the equality of sovereign states. Accordingly, customary international law rules on functional and personal immunities (discussed below) ‘have developed to ensure reciprocal respect among states for their sovereignty’; more precisely, they ‘aim at preventing states from interfering with the fulfilment of’ sovereign activities by foreign state representatives in their territories, and at preventing possible abuses by the

²³³ Rome Statute art 27 stipulates that the ‘Statute shall apply equally to all persons without any distinction based on official capacity’, that such capacity cannot ‘exempt a person from criminal responsibility’ under the Statute or be a basis for mitigation of sentence, and that ‘[i]mmunities or special procedural rules which may attach to the official capacity of a person, whether under national or international law, shall not bar the Court from exercising its jurisdiction over such a person’. The Pre-Trial Chamber II, in an obiter dictum, confirmed the scope of art 27(2) of the Statute, stating that ‘the Statute cannot impose obligations on third States without their consent’ and ‘[t]hus, the exception to the exercise of the Court’s jurisdiction provided in article 27(2) of the Statute should, in principle, be confined to those States Parties who have accepted it’. See ICC-02/05-01/09-195 (note 184 above) at para 26. Also highlighted in *al-Bashir* (note 19 above) at para 49.

²³⁴ The position in the ICC has also been the position in other international tribunals. See, eg, Updated Statute of the International Criminal Tribunal for the Former Yugoslavia (2009) art 7(2); Statute of the International Criminal Tribunal for Rwanda (1994) art 6(2); and Statute of the Special Court of Sierra Leone (2000) art 6(2). It has thus allowed for warrant of arrests to be used for sitting heads of states such as al-Bashir. Also, a warrant of arrest was issued against Charles Taylor while he was still President of Liberia with the Special Court of Sierra Leone subsequently confirming that an international court is not barred from prosecuting a head of state. See *Prosecutor v Charles Taylor (Decision on Immunity from Jurisdiction)* SCSL-2003-01-AR72(E) A Ch (31 May 2004) at paras 51–53. Further, Hissène Habré, as a former head of state, could also not rely on immunity. See African Union Report of the Committee of Eminent African Jurists on the Case of Hissène Habré (2006) at paras 12–14, available at http://www.peacepalacelibrary.nl/ebooks/files/habreCEJA_Report0506.pdf; Statute of the Extraordinary African Chambers within the Courts of Senegal created to Prosecute International Crimes committed in Chad between 7 June 1982 and 1 December 1990, at 10(3), available at <https://www.hrw.org/news/2013/09/02/statute-extraordinary-african-chambers>; and P Gaeta ‘Ratione Materiae Immunities of Former Heads of State and International Crimes: The Hissène Habré Case’ (2003) 1 *Journal of International Criminal Justice* 186–196. Note that the Extraordinary African Chambers within the courts of Senegal (EAC) is mentioned here because, ‘while the EAC has only a few international aspects, it is still properly considered an internationalized criminal tribunal, although it is at the limit of this category’. S Williams ‘The Extraordinary African Chambers in the Senegalese Courts: An African Solution to an African Problem?’ (2013) 11 *Journal of International Criminal Justice* 1139, 1140.

²³⁵ See Schabas (note 61 above) at 247 (Rightly argues that arts 27 and 98(1) of the Rome Statute are not inconsistent or incompatible).

territorial state of its powers and authorities.²³⁶ Criminal jurisdiction exercised by international courts is lacking this 'very *rationale*'.²³⁷ The ICC, for example, is not an organ of a particular state. Its mandate to prosecute the most serious crimes derives from the international community as a whole.²³⁸ Therefore, the ICC's judicial activity is neither 'an expression of the sovereign authority of a state upon that of another state' nor 'a form of "unduly" interfering with the sovereign prerogatives of another state'.²³⁹ As a result, the rules of customary international law on sovereign immunity would not apply when international courts exercise criminal jurisdiction. While this is generally accepted for functional immunity,²⁴⁰ in relation to personal immunity this is controversial (thus, international tribunals have explicitly excluded its application before them or ruled against its applicability based on the international nature of the tribunals).²⁴¹

Pursuant to art 27 of the Rome Statute, states parties have agreed not to invoke immunities based on official capacity of a person in ICC proceedings.²⁴² While in relation to *al-Bashir*, Sudan is not a party to the Rome Statute (and the UNSC does not render it a state party),²⁴³ the implication of the UNSC referral is that

²³⁶ See Gaeta (note 172 above) at 320, with reference to personal immunities.

²³⁷ Ibid.

²³⁸ See Rome Statute preamble. See also Gaeta (note 171 above) at 321; and *Taylor* (note 234 above) at para 51.

²³⁹ Gaeta (note 171 above) at 321.

²⁴⁰ Confirmed in *Arrest Warrant* (note 50 above) at para 61. See also *Blaškić* ICTY Appeals Chamber (24 October 1997) at para 41 (confirming that in international law functional immunity does not apply for crimes against humanity, genocide and war crimes); Cryer, Friman, Robinson & Wilmschurst (note 1 above) at 545.

²⁴¹ Against the general applicability of customary international law rules on personal immunity in international courts, see, eg, Van der Vyver (note 172 above) at 559–579 and 570–573; Gaeta (note 171 above) at 320 and 322 (stating that article 27 merely 'restates an already existing principle concerning the exercise of jurisdiction by *any* international criminal court'); P Gaeta 'Official Capacity and Immunities' in A Casesse, P Gaeta & JRWD Jones (eds) *The Rome Statute of the International Criminal Court: A Commentary* (2002) 975, 991; *Taylor* (note 234 above). Assuming its general existence in international courts, see, eg, *Arrest Warrant* (note 50 above) at para 61 (stating that an official can be tried when personal immunity falls away or if the statute of a specific tribunal excludes it); ICC-02/05-01/09-242 (note 157 above) at para 43 (calling article 27 'an exception to Head of State immunity when international courts seek a Head of State's arrest for the commission of international crimes'); *al-Bashir* (note 19 above) at paras 59 and 78 (stating that article 27 constitutes a waiver of immunity that 'their nationals would otherwise have enjoyed under customary international law'). For further discussion on this, see CD Akande 'International Law Immunities and the International Criminal Court' (2004) 98 *American Journal of International Law* 415–419; and Bantekas (note 1 above) at 133. See also Cryer, Friman, Robinson & Wilmschurst (note 1 above) at 549–550 (Noting that state consent is required to waive personal immunity).

²⁴² Acknowledged in *al-Bashir* (note 19 above) at paras 59 and 78 (States parties have waived rights to immunity that their nationals would have enjoyed).

²⁴³ See Tladi (note 234 above) at 1043. But see Akande (note 198 above) at 341–342 (Arguing that the UNSC resolution rendered Sudan akin to an ICC State Party. However, in light of the general principle of international law *pacta tertiis nec nocent nec prosunt*, enshrined in art 34 of the Vienna Convention of the Law of Treaties, this is at the very least problematic).

the investigation and prosecution will take place in accordance with the Statute, Elements of Crime and Rules of Procedure and Evidence.²⁴⁴

Hence, the ICC Pre-Trial Chamber I, in justifying its ability to exercise jurisdiction, considered ‘that the current position of Omar al-Bashir as Head of a state which is not a party to the Statute, has no effect on the Court’s jurisdiction over the present case’.²⁴⁵ This was based, *inter alia*, on the following considerations: firstly, the core principles of art 27 (basically a recital of the provision with further explanation) and, secondly, the implication of the referral being that the investigation and prosecution will take place in accordance with the Statute, Elements of Crime and Rules of Procedure and Evidence.²⁴⁶ These considerations suggest that ‘the Security Council has implicitly adopted Article 27 and thus implicitly sanctioned the exercise of jurisdiction by the Court over a serving head of state who would otherwise be immune from jurisdiction’.²⁴⁷ The ICC Pre-Trial Chamber I has thus subsequently argued that the UNSC implicitly waived al-Bashir’s personal immunity *in the ICC proceedings* by referring the situation in Darfur to the ICC while acting under Chapter VII of the UN Charter.²⁴⁸ This line of reasoning is mainly based on the fact that UN member states, and therefore also Sudan, are required to carry out Chapter VII measures by virtue of art 25 of the UN Charter. That is supported by art 103 of the UN Charter which ensures that, in the event of a conflict, obligations under the UN Charter prevail over *all* obligations ‘under any other international agreements’. Further, the ICC’s view is that a UNSC referral resolution requiring ‘full’ cooperation from a UN member state who is a non-state party to the Rome Statute would be rendered meaningless if it had to be interpreted to exclude an implicit waiver of immunities.²⁴⁹

A consideration of the Genocide Convention presents an alternative argument in relation to al-Bashir’s immunity in the context of an arrest and surrender to the ICC. Pursuant to art IV of the Convention, official capacity cannot be raised as a defence to a prosecution for genocide. It provides that ‘[p]ersons committing genocide or any of the other acts enumerated in article III shall be punished, whether they are constitutionally responsible rulers, public officials or private

²⁴⁴ It is worth noting that the UNSC could indeed seek to impose treaty obligations on non-state parties while acting under Chapter VII. See, eg, UNSC Resolution 1373 (2001)(Imposing obligations on all States arising from the Convention for the Suppression of the Financing of Terrorism (1999)); UNSC Resolution 1874 (2009)(Imposing on North Korea the NPT treaty after it had announced its withdrawal); UNSC Resolution 1757 (2007)(Giving effect to an agreement between Lebanon and the UN to create the Special Tribunal for Lebanon after the Lebanese parliament refused to ratify it).

²⁴⁵ ICC-02/05-01/09-3 (note 197 above) at para 41.

²⁴⁶ *Ibid* at paras 42–45.

²⁴⁷ Akande (note 198 above) at 336; and Schabas (note 61 above) at 246.

²⁴⁸ See ICC-02/05-01/09-242 (note 158 above) at para 7, citing ICC-02/05-01/09-195 (note 184 above) at paras 29 and 31. On extending his implicit waiver to the horizontal level, see Part IV.A.2 below.

²⁴⁹ See ICC-02/05-01/09-195 (note 184 above) at para 29. In relation to states parties to the Rome Statute, the ICC is of the view that ‘[t]o interpret article 98(1) [of the Statute] in such a way so as to justify not surrendering Omar al-Bashir on immunity grounds would disable the Court and international criminal justice in ways completely contrary to the purpose of the Statute [the concerned state party] has ratified’. *Prosecutor v Omar Hassan Ahmad al-Bashir (Decision Pursuant to Article 87(7) of the Rome Statute on the Failure by the Republic of Malawi to Comply with the Cooperation Requests issued by the Court with Respect to the Arrest and Surrender of Omar Hassan Ahmad al-Bashir)* ICC-02/05-01/09-139 (12 December 2011) PT Ch I (‘*Malawi Decision*’) at para 41.

individuals'. Akande argues that 'the provision must also be taken as removing any procedural immunities, as the availability of any such immunities would mean that the persons mentioned in Article IV are not punished'.²⁵⁰ Should this interpretation be accepted, then pursuant to the ICJ's ruling on obligations to cooperate with a competent international tribunal,²⁵¹ and considering that Sudan is a party to the Genocide Convention plus al-Bashir is wanted for genocide charges, states parties to the Rome Statute that are also parties to the Convention can arrest al-Bashir without any concerns regarding immunities. '[T]he obligation of ICC parties to arrest is based on the acceptance of the ICC's jurisdiction by that party and the imposition of ICC jurisdiction on Sudan'; and 'the removal of immunity is based on the acceptance of the Genocide Convention by the arresting party and by Sudan'.²⁵²

In any event, art 27 of the Rome Statute (also art IV read with art VI of the Genocide Convention) refers to the disregard of immunities *when the ICC exercises its jurisdiction*. This does not exhaust the immunity question in *al-Bashir* since al-Bashir is not in the ICC's custody, leading to the question whether or not he is immune from arrest by national authorities cooperating with the ICC.²⁵³ Put differently, *al-Bashir* relates to the *exercise of criminal jurisdiction by domestic authorities* of foreign states, more precisely in our case South African authorities.

2 *International Law in National Criminal Proceedings in Foreign States*

As noted previously, the rules of customary international law on functional and personal immunities involve guarantees for certain government officials vis-à-vis the domestic authorities of the foreign state.²⁵⁴ They enjoy immunity on the basis of the governmental conduct or functions that they carry out (functional immunity) or on the basis of their status such as head of state and diplomats (personal immunity).²⁵⁵ The former only covers specific conduct on behalf of a state and therefore does not provide complete immunity, but immunity for that conduct does not fall away when the person's official role comes to an end. The latter provides absolute immunity for all actions, but only for the duration that the person holds their representative status.²⁵⁶ While there is an exception in relation

²⁵⁰ Akande (note 198 above) at 350.

²⁵¹ See Part IV.C.1.c.

²⁵² Akande (note 198 above) at 351.

²⁵³ See *al-Bashir* (note 19 above) at paras 76–77. But see Van der Vyver (note 172 above) at 573 (Without distinguishing the question of immunities before international and national courts, states: 'If President al-Bashir does not enjoy sovereign immunity for purposes of the ICC, there is no immunity that needs to be waived.')

²⁵⁴ This has been confirmed in, eg, *Arrest Warrant* (note 50 above) and *Jurisdictional Immunities* (note 227 above).

²⁵⁵ See Cryer, Friman, Robinson & Wilmshurst (note 1 above) at 533; and R Slye 'Immunities and Amnesties' in M du Plessis (ed) *African Guide to International Criminal Justice* (2008) 182.

²⁵⁶ See Cryer, Friman, Robinson & Wilmshurst (note 1 above) at 533–534; and Van Schaack & Slye (note 2 above) at 971.

to functional immunity when it comes to serious international crimes,²⁵⁷ personal immunity applies absolutely to the government official in question during his/her time in office – no action can be taken by domestic authorities of foreign states and before domestic foreign courts even in relation to serious international crimes. That rule applies (also) when it comes to the arrest and surrender of al-Bashir to the ICC by the competent domestic authorities of a foreign state.²⁵⁸

States parties to the Rome Statute arguably have, by virtue of art 98, agreed to derogate from customary international law on immunities in relation to Rome Statute crimes when their domestic authorities exercise jurisdiction regarding the arrest of a person from a state party for surrender to the ICC or to another state party.²⁵⁹ With regard to the surrender of a person (who ordinarily enjoys immunity) from a non-state party to the ICC, the ICC arguably²⁶⁰ has no permission to proceed with a request for surrender if it ‘would require the requested State to act inconsistently with its obligations under international law with respect to the State or diplomatic immunity of a person ... of a third State’.²⁶¹ This implies that in national proceedings in foreign states in relation to the surrender of a person representing a state party, there is such permission and thus surrender would be consistent with state obligations under customary international law on immunities. Due to the *pacta tertiis* principle, this (derogation) could only be done among states parties. Therefore, generally speaking, South African domestic authorities continue to be bound by the rules of customary international law on personal immunities when it comes to the arrest and the surrender of individuals from a non-state party to the Rome Statute, such as Sudan. In particular, and as Gaeta puts it,

²⁵⁷ To commit serious human rights or *jus cogens* violations is not recognised as one of the functions of statehood and can therefore not be attributed to the state, ie lays outside available sovereign prerogatives. See, eg, Bantekas (note 1 above) at 129 et seq; Cryer, Friman, Robinson & Wilmschurst (note 1 above) at 542–545; Van Schaack & Slye (note 2 above) at 968–974 and 976; and *Arrest Warrant* (note 50 above) at paras 47–55. Regarding the example in *R v Bow Street Stipendiary Magistrate and Others, ex parte Pinochet Ugarte (No. 3)* [1999] 2 All E.R. 97 where torture was not seen as a state function, see R van Alebeek ‘The Pinochet Case: International Human Rights on Trial (2000) 71 *British Yearbook of International Law* 29; A Bianchi ‘Immunity versus Human Rights: The Pinochet Case’ (1999) 10 *European Journal of International Law* 237; and A Pillay ‘Revisiting Pinochet: The Development of Customary International Criminal Law (2001) 17 *South African Journal of Human Rights* 477.

²⁵⁸ Acknowledged in *al-Bashir* (note 19 above) at paras 67–84, esp 73 and 84; and in *SALC* (note 23 above) at para 46, fn 50. See also ICC-02/05-01/09-195 (note 184 above) at para 25 (‘At the outset, the Chamber wishes to make it clear that it is not disputed that under international law a sitting Head of State enjoys personal immunities from criminal jurisdiction and inviolability before national Courts of foreign states even when suspected of having committed one or more of the crimes that fall within the jurisdiction of the Court.’)

²⁵⁹ But see JK Kleffner ‘The Impact of Complementarity on National Implementation of Substantive International Criminal Law’ (2003) 1 *Journal of International Criminal Justice* 86, 103–106 (Stating that ‘article 98(1) only deals with cases related to the exercise of jurisdiction by the ICC in relation to the arrest for surrender. Thus domestic authorities of foreign states continue to be bound by the rules of customary international law on personal immunities when it comes to the need to surrender those individuals to the ICC’ (emphasis added).)

²⁶⁰ As indicated in Part II.C.1.a above, art 98(1) could be interpreted as a ‘discretion’ but could also be interpreted as a ‘prohibition’.

²⁶¹ Rome Statute art 98(1) (emphasis added).

[t]here is no rule in international law that allows states parties to lawfully disregard those immunities to comply with a request for surrender by an international criminal court. The fact that the ICC ... is endowed with jurisdiction over a particular case but is deprived of enforcement powers, does not imply that national judicial authorities are permitted to do whatever an international court asked them to do.²⁶²

Hence, the ICC Pre-Trial Chamber I has subsequently argued that “there also exists no impediment at the horizontal level” regarding the arrest and surrender to the Court of Omar al-Bashir.²⁶³ It extended the assertion of an implicit waiver of his personal immunity in the ICC proceedings by implicitly adopting art 27 of the Rome Statute (also) to national proceedings in foreign states.²⁶⁴ The main arguments for such an extension are identical to the arguments in support of an implicit waiver mentioned above.²⁶⁵

Arguments against a waiver implied by the referral could be found in, firstly, art 98(1) of the Rome Statute and, secondly, in the legal effect of a UNSC referral within the purposes of the Rome Statute. To begin with, art 98(1) of the Rome Statute is the only provision in the treaty which makes provision for the possibility of waiving immunity, and the waiver has to be given by the third state – the non-party state that the person claiming immunity represents.²⁶⁶ In the case of al-Bashir, the third state is Sudan (which has not waived al-Bashir's immunity). No alternatives are provided for in the Statute. In addition, the legal effect of a UNSC referral is provided for in art 13(b) of the Rome Statute²⁶⁷ – a referral serves as a trigger for the ICC's jurisdiction and this includes the jurisdiction over crimes committed in the territory or by nationals of a non-state party to the Rome Statute.²⁶⁸ Had the parties to the Rome Statute intended to confer further legal effects to UNSC referrals (other than triggering jurisdiction), they could and should have explicitly stated so. Thus, within the purposes of the Rome Statute and the UNSC referral provision in the Statute, an implied waiver possibility in national proceedings by virtue of a UNSC referral is, at the very least, problematic.²⁶⁹

²⁶² Gaeta (note 171 above) at 325 and 328. See also M Blommesteijn & C Ryngaert ‘Exploring the Obligations for States to Act upon the ICC's Arrest Warrant for Omar Al-Bashir: A Legal Conflict between the Duty to Arrest and the Customary Status of Head of State Immunity’ (2010) 6 *Zeitschrift für Internationale Strafrechtsdogmatik* 428, 438–440. See also CD Akande ‘The Effect of Security Council Resolutions and Domestic Proceedings on State Obligations to Cooperate with the ICC’ (2012) 10 *Journal of International Criminal Justice* 299, 301; and *Malawi Decision* (note 249 above) at para 15.

²⁶³ See ICC-02/05-01/09-242 (note 157 above) at para 7, citing ICC-02/05-01/09-195 (note 184 above) at paras 29 and 31.

²⁶⁴ *Ibid.*

²⁶⁵ See above Part V.A.1.

²⁶⁶ Rome Statute art 98(1) stipulates: ‘... unless the Court can first obtain the cooperation of that third State for the waiver of the immunity.’

²⁶⁷ The UNSC is mentioned too in Rome Statute art 87(7) (‘the ICC may refer a matter to the UNSC if a party fails to cooperate, where the case was referred to the Court by the UNSC’).

²⁶⁸ Rome Statute art 13 reads, in relevant part: ‘The Court may exercise its jurisdiction with respect to a crime referred to in article 5 in accordance with the provisions of this Statute if: ... (b) A situation in which one or more of such crimes appears to have been committed is referred to the Prosecutor by the Security Council acting under Chapter VII of the Charter of the United Nations’.

²⁶⁹ See also Gaeta (note 171 above) at 324.

However, the ICC Pre-Trial Chamber II has held that ‘the cooperation envisaged ... was meant to eliminate any impediment to the proceedings before the Court, including the lifting of immunities’.²⁷⁰ A different interpretation, it held, would render the UNSC decision ‘senseless ... since immunities attached to al-Bashir are procedural bars from prosecution before the Court’.²⁷¹ Thus, “cooperation of ... [Sudan] for the waiver of the immunity” as required under the last sentence of article 98(1) of the Statute, was already ensured by the language used in paragraph 2 of SC Resolution 1593 (2005).²⁷²

Assuming that the UNSC is permitted to waive al-Bashir’s immunity, in Resolution 1593 (2005) the UNSC *merely urges* all states to cooperate, without stating that this cooperation implied waiver of immunity for it to be meaningful, while the obligation to ‘fully cooperate’ is placed on Sudan and other parties to the conflict.²⁷³ The UNSC’s general encouragement neither calls for states parties to disregard customary international law rules on personal immunities for purposes of cooperation with the ICC nor can it be construed as implying that states parties are authorised to violate these rules without bearing any international responsibility.²⁷⁴ Had the UNSC intended this, it could and should have explicitly said so, especially since Sudan, as a UN Member, would then indeed have to accept such a Chapter VII decision by virtue of art 25 of the UN Charter. It is worth noting that such an interpretation does not, as claimed by the ICC Pre-Trial Chamber I,²⁷⁵ render the UNSC resolution meaningless (at least in theory) since the resolution requires cooperation from some states, including Sudan, and in the ICC proceeding al-Bashir has no immunity.

In the absence of a waiver, to disregard President al-Bashir’s personal immunities in national proceedings in foreign states and surrender him to the ICC would constitute an international wrongful act, even though this wrongful act would not infringe upon the jurisdiction of the ICC over al-Bashir. Whether it would be a wrongful act in terms of domestic law in relation to al-Bashir will depend on the status that a state’s constitution accords to customary international law within its domestic legal system. We address that issue in the context of South Africa below.

B Domestic Law

In terms of domestic law, immunities and privileges in the South African context are regulated in the Diplomatic Immunities and Privileges Act (DIPA).²⁷⁶ Section 4(1) of the DIPA stipulates that a ‘head of state is immune from criminal ... jurisdiction of the courts of the Republic, and enjoys such privileges as ... heads

²⁷⁰ ICC-02/05-01/09-195 (note 184 above) at para 29.

²⁷¹ Ibid.

²⁷² Ibid.

²⁷³ See Parts III.C.1.b and IV.A.1 above.

²⁷⁴ See Gaeta (note 171 above) at 332.

²⁷⁵ See Part V.A.1 above, esp fn 259.

²⁷⁶ Diplomatic Immunities and Privileges Act 37 of 2001. This Act incorporates several UN conventions on diplomatic and consular privileges and immunities. For a discussion on DIPA, see, eg, H Strydom ‘Diplomatic privileges and immunities’ in H Strydom (ed) *International Law* (2015) ch 10, 307, 316.

of state enjoy in accordance with the rules of customary international law'. As a result (specifically in terms of the DIPA), customary international law on personal immunities should apply to al-Bashir.

With regards to additional immunity and privileges afforded at the domestic level, South Africa, as the host nation of the AU Summit, entered into a hosting agreement with the AU, in which it committed, under art VIII of the agreement, to grant immunities and privileges contained in

[s]ections C and D and Article V and VI of the General Convention of the Privileges and Immunities of the OAU [Organisation of African Unity] to the members of the Commission and Staff Members, delegates and other representatives of Inter-Governmental Organisations attending the Meetings.²⁷⁷

Pursuant to s 5(3) of the DIPA,²⁷⁸ the agreement was proclaimed in the *Government Gazette* by the Minister of International Relations and Cooperation²⁷⁹ – the agreement was thus incorporated into domestic law. On the basis of the hosting agreement and the Ministerial proclamation, the government argued that al-Bashir was entitled to immunities during the AU summit and two days after its conclusion, and could therefore not be arrested.²⁸⁰

However, the SCA rightly held that 'the hosting agreement did not confer any immunity on President al-Bashir and its proclamation by the Minister of International Relations and Cooperation did not serve to confer any immunity on him'.²⁸¹ Firstly, the proclamation under s 5(3) of the DIPA – the provision the government invoked – applies to organisations and their representatives. According to s 1(iv) of the DIPA, "organisation" means an intergovernmental organisation of which two or more states or governments are members and which the Minister has recognised for the purposes of this Act'. This refers to organisations such as the AU or African Commission on Human and Peoples' Rights and does not include member states or their representatives, such as heads of states.²⁸² The SCA therefore held that the hosting agreement provides immunity only for representatives and officials of the AU and organisations and not for those of states. It does not, therefore, provide immunity for heads of states and state delegates.²⁸³ In particular, the SCA held that, based on an analysis of art VIII of the hosting agreement and the description of the AU Assembly, heads of states attend the AU summit as 'the embodiment of the member state not delegates from them'.²⁸⁴ Secondly, even though additional immunity can be

²⁷⁷ *Al-Bashir* (note 19 above) at para 11.

²⁷⁸ DIPA s 5 reads: 'Any organisation recognised by the Minister for the purposes of this section and any official of such organisation enjoy such privileges and immunities as may be provided for any agreement entered into with such organisation or as may be conferred on them by virtue of section 7(2).'

²⁷⁹ See *al-Bashir* (note 19 above) at para 12.

²⁸⁰ *Ibid* at para 13.

²⁸¹ *Ibid* at para 47.

²⁸² *Ibid* at paras 41–42.

²⁸³ *Ibid* at paras 42 and 47.

²⁸⁴ *Ibid* at paras 44–46.

granted to heads of states through s 7 of the DIPA²⁸⁵ – also recognised in s 4(1)(a) of the Act²⁸⁶ – it would be problematic to do so in al-Bashir's case, as it would be in contravention of the ICC Act which explicitly prohibits head of state immunity for Rome Statute crimes.

The ICC Act follows a similar approach to the Rome Statute on the question of immunities and privileges of government officials. Section 4(2) of the ICC Act diverges from customary international law by excluding them from using immunity as a basis for a defence to a listed crime or for the reduction of their sentence following conviction before South African courts.²⁸⁷ These persons include head of states or governments.²⁸⁸ Section 10(9) of the ICC Act extends non-recognition of head of states immunity to an order to surrender them to the ICC. It declares that '[t]he fact that the person to be surrendered is a person contemplated in section 4(2)(a) or (b) does not constitute a ground for refusing to issue an order', and that a person 'be surrendered to the [ICC] and that he or she be committed to prison pending such surrender'.²⁸⁹

The government argued that s 4(2) of the ICC Act does not remove head of state immunity in the context of an arrest, and that s 10(9) applies only in the context of surrender.²⁹⁰ The SCA disagreed. It held that s 10(9) applies equally to an arrest for purposes of surrendering head of states to trial before the ICC.²⁹¹ Even though 's 10 of the ICC Act deals only with the surrender of persons who had already been arrested under s 9 and ... the latter section [is] silent on the question of immunity',²⁹² to interpret this as recognising immunity for arrests would, in Wallis JA's words –

creat[e] an absurdity. If it were correct, then any person entitled on any basis to claim immunity would challenge their arrest by way of an *interdict de libero homine exhibendo* ... and demand their release. So the only people who could be brought before a magistrate under s 10 would be those who had no grounds for claiming immunity. But then s 10(9) would serve no purpose at all. It would be entirely redundant, because there would be no possible situation in which a person brought before the magistrate under s 10(1) would be a person referred to in ss 4(2)(a) or (b). Needless to say such an interpretation is to be avoided.²⁹³

Moreover, '[t]he ordinary principle of interpretation is that conferral of powers conveys with it all ancillary powers necessary to achieve the purpose of that power'.²⁹⁴ The purpose of the power to surrender a person charged with international crimes to the ICC is to ensure that perpetrators of such crimes

²⁸⁵ Ibid at para 42. DIPA s 7 includes the granting of immunities and privileges to 'any person' (in addition to organisations) through notice in the Government Gazette.

²⁸⁶ DIPA s 4(1)(a) reads: 'A head of state is immune from the criminal and civil jurisdiction of the courts of the Republic, and enjoys such privileges as ... heads of state enjoy in accordance with the rules of customary international law.' Note that this is immunity from jurisdiction of South African courts not international courts.

²⁸⁷ But see Tladi (note 230 above).

²⁸⁸ See ICC Act s 4(2)(a).

²⁸⁹ ICC Act s 10(5).

²⁹⁰ See *al-Bashir* (note 19 above) at paras 50 and 101.

²⁹¹ Ibid at para 101.

²⁹² Ibid.

²⁹³ Ibid. See also Ventura (note 20 above).

²⁹⁴ *Al-Bashir* (note 19 above) at para 95.

do not go unpunished. In order to achieve that purpose, it is necessary for the magistrate to have the power to issue a warrant of arrest to bring such persons before the ICC. Such an approach, as the SCA points out, 'is consistent with the constitutional requirement that the [ICC Act] be construed in a way that gives effect to South Africa's international law obligations and the spirit, object and purpose of the Bill of Rights'.²⁹⁵

Thus, since there is no differentiation in the ICC Act between government officials from states parties and non-states parties to the Rome Statute, al-Bashir does *not* enjoy immunity from jurisdiction before South African courts in relation to the issuing of an arrest warrant and an order to surrender him to the ICC.

The above interpretation of ss 10(9) and 4(2) of the ICC Act is inconsistent with customary international law on personal immunities before foreign domestic courts. But their wording does not allow for an alternative interpretation. Section 233 of the Constitution – which requires legislation to be interpreted consistently with customary international law – is therefore inapplicable. Sections 10(9) and 4(2) expressly exclude immunity for government officials, including heads of state, based on their status before South African courts.

In addition, it is worth noting that while s 4(2) of the ICC Act paraphrased the provisions of art 27 of the Rome Statute, there is an essential difference between the scope of the two provisions. The former applies to immunities before South African courts, and the latter to immunities when the ICC exercises jurisdiction. In other words, the provisions involve different levels of proceedings. More specifically, there is no obligation for states parties to apply the same immunity rules to states not party to the Rome Statute that are applicable among contracting parties.²⁹⁶ In fact, art 98(1) of the Rome Statute acknowledges that states parties can be in conflict with customary rules on immunity in relation to third states. Why else would it provide for the possibility of waiver? This reflects 'the will of the drafters to avoid, to the greatest extent possible, the obligations of contracting states to cooperate with the Court from becoming incompatible with international obligations binding a state party *vis-à-vis* a state not party to the ICC Statute'.²⁹⁷

The conflicting positions on immunities under the DIPA and the ICC Act – on the one hand granting immunity from criminal jurisdiction, and on the other denying immunity for international crimes – can be resolved by two of the general principles on conflict resolution:²⁹⁸ *lex posterior derogat legi priori* and *lex specialis derogat legi generali*. Firstly, the ICC Act was adopted later (in 2002) than the DIPA (in 2001) and is therefore *lex posterior*. Secondly, international criminal law is a specific part of criminal law and thus *lex specialis* in this regard, as confirmed in *al-Bashir*. Accordingly, 'the DIPA is a general statute dealing with the subject of immunities and privileges enjoyed by various people, including heads of states.

²⁹⁵ Ibid.

²⁹⁶ But see *al-Bashir* (note 19 above) (Stating that not granting immunity before South African courts would result in not complying with the obligation to cooperate under the Rome Statute).

²⁹⁷ Prost & Schlunck (note 172 above).

²⁹⁸ See Part IV.C.3, esp fn 218 above.

The [ICC] Act is a specific Act dealing with South Africa's implementation of the Rome Statute. In that special area the Implementation Act must enjoy priority'.²⁹⁹

Finally, so far as international crimes are concerned, customary international law rules on personal immunities are not directly applicable. According to s 232 of the SA Constitution, 'customary international law is law in the Republic' if it is consistent with the Constitution and Acts of Parliament. Granting al-Bashir personal immunity before South African courts would conflict with s 4(2) of the ICC Act: the rule that head of the states do not enjoy personal immunity for international crimes.

This shows that not granting sovereign immunity in the *al-Bashir* and *SALC* judgments is consistent with South African domestic law. The competing international and domestic law principles and obligations in relation to al-Bashir's personal immunity operate on different levels. As highlighted above, domestic laws cannot be invoked as justification for a state's failure to comply with its international obligations.³⁰⁰ International obligations, in turn, are no basis to override domestic legal obligations before domestic courts.

VI CONFLICTS BETWEEN INTERNATIONAL OBLIGATIONS RELATING TO IMMUNITY AND TO THE PROSECUTION OF INTERNATIONAL CRIMES

Conflicts between international obligations relating to immunity and those relating to prosecution are not a concern in *SALC* since immunity rules are not applicable in relation to the investigation of international crimes. But the situation in *al-Bashir* is different. Provided that there is an obligation to cooperate,³⁰¹ this would directly conflict with the customary international law rule on personal immunity, 'the grant of which is now understood as an obligation under customary international law'.³⁰² South Africa could not simultaneously comply with both obligations. The arrest and surrender of al-Bashir would render the granting of immunity in South African domestic proceedings impossible. And granting him immunity in South African domestic proceedings would preclude the SAPS from cooperating with the ICC by arresting al-Bashir and surrendering him to the ICC.³⁰³

²⁹⁹ *Al-Bashir* (note 19 above) at para 102.

³⁰⁰ See Part V.C.3.b.

³⁰¹ On the discussion whether the ICC has discretion in relation to a request of assistance or not and whether the *lex posterior* rule is applicable to the conflict between South Africa's obligation towards the AU to not cooperate with the ICC and the obligation to cooperate (stemming from the Rome Statute) or not, see Parts IV.C.1.a and IV.C.3 above.

³⁰² Crawford (note 41 above) at 487, esp fn 4 (Including detailed evidence that 'the existence of this obligation is supported by ample authority').

³⁰³ This even constitutes a norm conflict in terms of the strict definition, according to which, a conflict between two rules arises only where a state bound by them 'cannot simultaneously comply with its obligations'. W Jenks 'The Conflict of Law-Making Treaties' (1953) 30 *British Yearbook of International Law* 401, 426 (emphasis added). According to a broad approach, '[t]here is a conflict between norms, one of which may be permissive, if in obeying or applying one norm, the other norm is necessarily or potentially violated'. E Vranes 'The Definition of "Norm Conflict" in International Law and Legal Theory' (2006) 2(17) *European Journal for International Law* 395, 418. See also, Koskenniemi (note 220 above) at para 25 (Koskenniemi – on behalf of the ILC – adopts 'a wide notion of conflict as a situation where two rules or principles suggest different ways of dealing with a problem').

Since no explicit provision is made for a hierarchy between treaty and customary law obligations, the question of priority is highly controversial.³⁰⁴ There are four general approaches. Some scholars argue that 'the arrangement of the sources in paras (1)(a) to (c) [of art 38 of the ICJ Statute] ... does reflect a common-sense approach to the ranking of the sources',³⁰⁵ which must be applied by the courts accordingly (first approach).³⁰⁶ This approach is (partly) reconcilable with the view that even though there is no formal hierarchy as between treaties, customs and general principles, 'in most instances treaties' are regarded 'as the primary source, while custom is the secondary source' (second approach).³⁰⁷ This is based on the assertion that treaties are the primary means of norm creating and the most reliable source for determining (a state's) consent; providing predictability and certainty.³⁰⁸ Other scholars contend that both treaties and customs enjoy the same normative superiority vis-à-vis general principles, judicial decisions and academic writings because both are founded on the consent of states, which emphasises the consensual basis of international law (third approach).³⁰⁹ Finally, some writers claim (in line with classical international law),³¹⁰ that the wording of art 38(1) of the ICJ-Statute does not indicate (the existence of) a hierarchy between treaties, customs and general principles, thus they enjoy equal ranking – unless the general rule in question is one of *jus cogens* or an obligation *erga omnes* (fourth approach).³¹¹ What all four approaches have in common is that general rules of *jus cogens* and obligations *erga omnes* enjoy the highest status.³¹²

³⁰⁴ See A Pellet 'Article 38' in A Zimmermann, C Tomuschat & K Oellers-Frahm (eds) *The International Court of Justice – A Commentary* (2006) 778. See generally H Villiger *Customary International Law and Treaties* (1985).

³⁰⁵ J Stug & TW Bennett *Introduction to International Law* (2013) 12. See also G Fitzmaurice 'Some Problems Regarding the Formal Sources of International Law' (1958) *Symbolae Verzijl* 153; and J Kammerhofer 'Uncertainty in the Formal Sources of International Law: Customary International Law and some of its Problems' (2004) 15 *European Journal of International Law* 523. But see Akehurst (note 223 above) at 274 (Concluding this from the fact that the words *en ordre successif* have been deleted by the Sub-Commission of the third Committee of the First Assembly of the League of Nations from the first draft of art 38).

³⁰⁶ The ICJ uses treaties, customs and general principles in successive order and 'has organized a kind of complementarity between them'. Pellet (note 304 above) at 773.

³⁰⁷ J Dugard *International Law: A South African Perspective* (4th Edition, 2011) 27. See also Koskenniemi (note 220 above) at 47 (para 85) ('Treaties generally enjoy priority over custom').

³⁰⁸ See, eg, L Le Fur 'Règles Générales du droit de la Paix' (1935) 54 *Recueil des Cours* 5, 212.

³⁰⁹ See *Case concerning Military and Paramilitary Activities in and against Nicaragua (Nicaragua v USA)* ICJ Reports (1986) 14 at para 175 (The ICJ confirmed that treaties and customs can be equal in ranking).

³¹⁰ See Dugard (note 307 above) at 25.

³¹¹ See, eg, P Dupuy *Droit International Public* (8th Edition, 2006) 370 et seq (Acknowledging that there is a hierarchy between these three sources and 'secondary source-setting processes envisaged in treaty rules'). Judicial decisions and writings (lit d), for example, clearly have a subordinate function within the hierarchy of sources in light of their description as 'subsidiary means for the determination of rules of law'. Cassese (note 217 above) at 154. See also C Thiele 'Fragmentierung des Völkerrechts als Herausforderung für die Staatengemeinschaft' (2008) 46 *Archiv des Völkerrechts* 1, 7 (Noting that the arrangement in art 38 of the ICJ Statute is merely a listing from the generally more specific to the more general rules).

³¹² See, eg, Cassese (note 217 above) at 155 and 199 et seq; Shaw (note 41 above) at 123 et seq; and AL Paulus 'Jus Cogens in a Time of Hegemony and Fragmentation' (2005) 3/4 (74) *Nordic Journal of International Law* 297.

If one follows the first or the second approach, ie the text of law has priority over any other source of law, regardless of whether the source is authoritative or substantive, then the treaty obligation to cooperate with the ICC has priority vis-à-vis the customary international law rule on personal immunity in national proceedings (*lex superior derogat legi inferiori* – a law higher in the hierarchy repeals the lower one).

If one follows the third or the fourth approach, ie treaty and customary law obligations are equal in ranking, conflict resolution between the obligation to cooperate with the ICC and to grant immunity in national proceedings is more complex.

To begin with, the conflict clause in art 103 of the UN Charter does not provide a solution. It applies between obligations deriving from the UN Charter and other treaty obligations.³¹³ Firstly, neither South Africa's obligation to cooperate with the ICC nor its obligation to grant immunity in national proceedings stems from the UN Charter. And, secondly, for the sake of argument, even if one assumes, as some scholars do,³¹⁴ that South Africa's obligation to cooperate would also derive from the UNSC referral, and thus the UN Charter, the obligation to grant immunity in national proceedings is not a treaty obligation, as required by art 103 of the UN Charter,³¹⁵ but one deriving from customary international law.

A further consideration is that the Rome Statute does not contain a conflict clause for the conflicting obligations in question.³¹⁶ That is, a clause that regulates 'the relation between the provisions of the treaty and those of another treaty [or any other international law rule] or of any other treaty relating to the matters with which the treaty deals'³¹⁷ and aims to clarify which provision prevails in case of conflict.³¹⁸ More precisely, the wording of art 98 of the Rome Statute reflects that

³¹³ UN Charter art 103 reads: 'In the event of a conflict between the obligations of the Members of the United Nations under the present Charter and their obligations *under any other international agreement*, their obligations under the present Charter shall prevail' (emphasis added).

³¹⁴ See, eg, Gaeta (note 171 above) at 326 et seq.

³¹⁵ The ICJ itself states (in the context of *jus cogens*): '[T]he relief which article 103 of the Charter may give the Security Council in case of conflict between one of its decisions and an operative *treaty obligation*'. (1993) *ICJ Reports* 325, 440; 95 *ILR* 43, 158 (emphasis added). But see Gaeta (note 171 above) at 326 (Seems to extend the wording of art 103, especially the phrase 'other international agreement' to all international obligations of UN members without further substantiation. This understanding is not covered by the explicit wording 'international agreement' which is used interchangeably in international law with the term 'treaty'. In terms of VCLT art 2(1)(a), "'treaty" means an international agreement concluded between States'). See also Koskenniemi (note 220 above) at paras 344–345 (After explaining the arguments for both sides, he suggest that art 103 of the UN Charter 'should be read extensively – so as to affirm that charter obligations prevail also over United Nations Member States' customary law obligations' based on, firstly, customs being usually more general than treaties and, secondly, on the fact that since UNSC decisions supposedly prevailing over customs, all other UN Charter obligations must too. However, customs can be more specific than treaties (see, eg, Akehurst (note 223 above) at 274–278; and Cassese (note 217 above)) and UNSC decisions are binding (UN Charter art 25) but do not automatically prevail over customary customs).

³¹⁶ It should be noted that Rome Statute art 21 does not apply in our present context as it is not per se a conflict clause and the hierarchical approach in it is restricted specifically to laws that the ICC has to apply.

³¹⁷ Koskenniemi (note 220 above) at 214.

³¹⁸ According to Prosper Weil this relative normativity is a 'threat to the integrity of international law'. P Weil 'Towards Relative Normativity in International Law?' (1983) 3(77) *American Journal of International Law* 413.

states parties want to ensure that their cooperation obligations with the ICC do not become incompatible with customary international laws on immunity that bind a state party to a non-state party.³¹⁹ But instead of clarifying the relation between the relevant obligations, states parties agreed merely that the ICC shall not proceed with its request for cooperation.³²⁰

This raises the question of whether the three above mentioned general principles on conflict resolution can be of assistance:³²¹ (1) *lex posterior derogat legi priori*; (2) *lex posterior generalis non derogat priori speciali*; or (3) *lex specialis derogat legi generali*. 'For the *lex specialis* principle to apply it is not enough that the same subject matter is dealt with by two provisions',³²² the same subject matter being the arrest and surrender of al-Bashir by the South African government.³²³ '[T]here must be ... a discernible intention that one provision is to exclude the other.'³²⁴ Neither the formal request by the Pre-Trial Chamber I to cooperate with the ICC in relation to al-Bashir's arrest and surrender nor the customary international law obligation to grant him personal immunity in national proceedings is more general vis-à-vis the other. Each obligation protects different legal interests: on the one hand, the prosecution of international crimes and on the other 'preventing states from interfering with the fulfilment of' sovereign activities by foreign state representatives in their territories, and preventing abuses by the territorial state of its powers and authorities. Under such circumstances, the *lex specialis* rule provides no adequate solution³²⁵ since both obligations are, to some extent, *lex specialis*.

Moreover, while South Africa's obligation to cooperate with the ICC in the arrest and surrender of al-Bashir exists since March 2009 and would thus be the later law in relation to the obligation to grant immunity in national proceedings, to apply the *lex posterior* rule between treaties and customs would ignore the fact that '[t]here is a presumption of interpretation ... that treaties are not intended to derogate from customary law, just as statutes in English law are presumed not to derogate from the common law'.³²⁶

VII CONCLUSION

The South African government has an obligation, both under international and domestic law, to investigate and prosecute international crimes as well as to cooperate in their investigation and prosecution. It has put in place a progressive legal framework for the enforcement of international criminal law within the country, in line with its obligations. In practice, however, the extent of the government's support for ending impunity for international crimes and for international criminal justice remains questionable. In *SALC*, the government was not willing to initiate an investigation into torture as a crime against humanity committed by Zimbabwean officials; while in *al-Bashir*, the government was not

³¹⁹ See Prost & Schlunck (note 172 above).

³²⁰ See Part III.C.1.a above.

³²¹ See note 218 above.

³²² Akehurst (note 223 above).

³²³ Otherwise there would be no conflict. See Koskeniemi (note 220 above) at para 21 et seq.

³²⁴ Akehurst (note 223 above).

³²⁵ See Thiele (note 311 above) at 8.

³²⁶ Akehurst (note 223 above) at 275 et seq.

willing to arrest al-Bashir for surrender to the ICC. The courts found the position of the government in both cases to be inconsistent with its international and domestic obligations.

The South African government has thus been severely criticised for its failure to comply with its international and domestic commitments in the *SALC* and *al-Bashir* debacle. For example, in *SALC*, the CC held that ‘SAPS’s decision not to conduct an investigation was wrong in law’ and that South Africa ‘cannot be seen to be tolerant of impunity for alleged torturers’ and it ‘dare not be a safe haven for those who commit crimes against humanity’.³²⁷ In relation to *al-Bashir*, the government of Botswana, for example, on 14 August 2015, while calling on all members of the ICC to cooperate with the court, condemned the South African government’s failure to arrest al-Bashir in the following words: ‘We therefore find it disappointing that President al-Bashir avoided arrest when he cut short his visit and fled, in fear of arrest, to his country’.³²⁸

There is no doubt that justice for the horrendous crimes committed in Zimbabwe and Sudan must be done, and that even those in power should be brought to justice. Nonetheless, in its efforts to ensure the effective prosecution of such crimes, the South African government has to balance competing international and domestic obligations, particularly its obligations to investigate and to cooperate in the investigation and prosecution of international crimes and its obligations in relation to immunity. The question of competing obligations was at issue in *al-Bashir* and could occur in *SALC* at a later stage, as long as the ICC Act is in force.³²⁹

In relation to *al-Bashir*, the South African government is ‘between a rock and a hard place’. On the one hand, and from an international law perspective, the South African government (based on an argument that al-Bashir’s immunity had not been waived)³³⁰ would have committed an international wrongful act had its police forces arrested and surrendered al-Bashir. The rules on customary international law on personal immunities apply between non-states parties to the Rome Statute and states parties in relation to national proceedings in foreign states (noting of course that this wrongful act will not infringe upon jurisdiction of the ICC over al-Bashir³³¹ and, from a South African domestic law perspective,

³²⁷ *SALC* (note 23 above) at paras 80–81.

³²⁸ Z Mogopodi ‘Botswana critical of SA and al-Bashir’ (17 June 2015) *IOL News*, available at www.iol.co.za/news/africa/botswana-critical-of-sa-over-bashir-1.1872656.

³²⁹ Initially, the government intended to repeal the ICC Act (see Implementation of the Rome Statute of the International Criminal Court Act Repeal Bill, [B 23 – 2016]) but did not follow through. See T Gqirana, ‘Rome Statute repeal bill withdrawn from Parliament’, *News24* (14 March 2017), available at <http://www.news24.com/SouthAfrica/News/rome-statute-repeal-bill-withdrawn-from-parliament-20170314>.

³³⁰ It should of course be acknowledged that if Sudan decides to cooperate with the ICC, then it would imply a relinquishment (by Sudan) of immunities that al-Bashir would have been entitled to, as Sudan would not be able to cooperate while asserting immunity at the same time.

³³¹ At least if one applies the principle *male captus bene detentus*, supported by state practice in 4 cases (*Attorney-General v Eichmann* (1961) 36 *International Law Reports* 5; *Frisbie v Collins* (1952) 342 US 519; *United States ex el Lujan v Gengler* (1975) 510 F.2d 62; *United States v Alvarez-Machain* (1992) 504 US 655) and rejected in 5 cases (*State v Ebrahim* 21 *International Law Materials* 888; *United States v Toscanino* (1974) 500 F.2d 267; *United States v Verdugo-Urquidez* (1990) 939 F.2d 1341; *Connelly v Director of Public Prosecutions* (1964) A.C. 1254; and *Bennett v Horseferry Road Magistrates’ Court* (1993) 3 All E.R. 138).

it will not be a wrongful act). On the other hand, the South African government (based on an interpretation that points to non-violation of art 98(1) of the Rome Statute and that disregards the obligation towards the AU) has committed an international wrongful act in not cooperating with the ICC in the arrest of al-Bashir, flouting its cooperation obligations under the Rome Statute as well as domestic law. Thus, and different to the statement in *al-Bashir*, it is not the 'relationship between the [ICC] Act and the head of state immunity conferred by customary international law ...[that] lies at the heart of this case'³³² but the hierarchy between an international treaty obligation and an obligation deriving from customary international law.

This unpleasant result remains relevant for two reasons. Firstly, the arrest warrant continues to constitute the legal basis upon which the South African government can surrender al-Bashir once he no longer enjoys immunity, for example, because he has ceased to be president or the South African government has obtained a waiver of immunity from Sudan. Secondly, as the SCA held, despite leaving the country, the order remains in force and can be enforced whenever al-Bashir visits South Africa.³³³ This has affected and will continue to affect the conduct of the South African government's diplomatic relations with Sudan.³³⁴

As *SALC* relates to investigations, the unpleasant result in *al-Bashir* is not at issue. Although, an investigation in itself involving Zimbabwean senior government officials could result in a strain in diplomatic relations between South Africa and Zimbabwe.³³⁵ However, should the investigations lead to the need to prosecute, then the South African government could, should any of the Zimbabwean senior government officials enjoy personal immunity, find itself also in a situation of balancing its obligation to prosecute with international obligations relating to immunities as well as other diplomatic and/or political considerations. In addition, an investigation is possible from within South African territory but it will face real challenges in obtaining information especially if Zimbabwe refuses to cooperate.

Finally, it should be noted that the ICJ has thus far not taken a clear stand on the hierarchy between an international treaty provision and a rule of customary international law, or on the application of conflict resolution principles in case of conflicts between sources of equal ranking. It has also not been asked to clarify the disputed question of whether a UNSC referral can be seen as a waiver of immunity before national courts.³³⁶ Against this background, and taking the *SALC* and *al-Bashir* debacles seriously, it is perhaps time for South Africa, Kenya or other concerned African governments to initiate and seek clarification on these issues at the ICJ, rather than speaking of withdrawal from the ICC. After all, withdrawal would not absolve them of their cooperation obligations in relation to

³³² *Al-Bashir* (note 19 above) at para 61.

³³³ Ibid at para 20 ('SALC indicated that any attempt by President al-Bashir to return to this country would prompt it to seek its enforcement').

³³⁴ On examples that the government has taken the order into account, *ibid*.

³³⁵ See Associated Foreign Press 'SA court ruling may strain Harare relations' (30 October 2014), available at <http://www.iol.co.za/news/africa/sa-court-ruling-may-strain-harare-relations-1772886>.

³³⁶ The SCA found it unnecessary to deal with the waiver of immunity questions by virtue of the UNSC referral based on conclusion regarding the ICC Act. *Al-Bashir* (note 19 above) at para 106.

al-Bashir.³³⁷ Also, if a withdrawal is not procedurally rational, it could be subject to judicial scrutiny as has been the case with South Africa. The North Gauteng High Court found that South Africa's decision to withdraw was unconstitutional and invalid, mainly because the Minister acted without prior parliamentary approval. The court ordered the government to revoke the notice.³³⁸ The court further clarified that, even if the withdrawal notice is given effect to at the international level (which 'does not take effect until a year'),³³⁹ 'domestically, [the] government would be obliged, among others, to arrest and surrender the indicted leaders, as long as the [ICC Act] is in force'.³⁴⁰ And, even in the absence of the Rome Statute and the ICC Act, South Africa would still have obligations to investigate and/or prosecute acts of torture under the Torture Act and the GCA.

POSTSCRIPT

On 6 July 2017, the ICC Pre-Trial Chamber II issued its decision on South Africa's non-compliance with the ICC's request to arrest and surrender al-Bashir.³⁴¹ It unanimously found South Africa to have 'failed to comply with its obligations under the Statute by not executing the Court's request ... while [al-Bashir] was on South African territory between 13 and 15 June 2015'.³⁴² This finding was based on a number of considerations, including but not limited to the following: Firstly, by virtue of the UNSC referral, while acting under Chapter VII of the UN Charter, 'the rights and obligations as provided for in the Statute, including article 27(2), are applicable to Sudan'. The Chamber asserts that this renders al-Bashir's immunity under customary international law (ie makes the rule) inapplicable vis-à-vis states parties to the Rome Statute when required to cooperate in his arrest and surrender.³⁴³ Secondly, 'article 98 of the Statute – even if applicable to the present situation – does not foresee the possibility for a requested State Party to unilaterally refuse compliance with a Court's request for arrest and surrender'.³⁴⁴ Thirdly, the host agreement between South Africa and the AU did not grant immunity to heads of states attending the AU summit.³⁴⁵ Fourthly, despite South Africa's interactions with the ICC between 11 and 13 June 2015, it was still under

³³⁷ See Rome Statute art 127(2) which reads: 'A State shall not be discharged, by reason of its withdrawal, from the obligations arising from this Statute while it was a Party to the Statute, including any financial obligations which may have accrued. Its withdrawal shall not affect any cooperation with the Court in connection with criminal investigations and proceedings in relation to which the withdrawing State had a duty to cooperate and which were commenced prior to the date on which the withdrawal became effective, nor shall it prejudice in any way the continued consideration of any matter which was already under consideration by the Court prior to the date on which the withdrawal became effective.' Confirmed in Memorandum on the Objects of the Implementation of the Rome Statute of the International Criminal Court Act Repeal Bill 2016 (2016) at para 1.6.

³³⁸ *Democratic Alliance* (note 12 above) at para 84.

³³⁹ *Ibid* at para 47, referring to Rome Statute art 127(1).

³⁴⁰ *Democratic Alliance* (note 12 above) at para 66.

³⁴¹ *The Prosecutor v Omar Hassan Ahmad Al Bashir (Decision under article 87(7) of the Rome Statute on the non-compliance by South Africa with the request by the Court for the arrest and surrender of Omar al-Bashir)* ICC-02/05-01/09-302 (6 July 2017) PT Ch II.

³⁴² *Ibid* at paras 123 and 140.

³⁴³ *Ibid* at para 107.

³⁴⁴ *Ibid* at para 108.

³⁴⁵ *Ibid* at para 133.

a duty to arrest al-Bashir while he was in the country and surrender him to the court.³⁴⁶ Finally, despite its finding, the Chamber did not deem a referral of the matter to the UNSC or the Assembly of States Parties necessary, considering that this is a discretionary power of the Chamber, in addition to the significance of South Africa being the first state party to seek to consult with the ICC under art 97 of the Statute, including seeking from the ICC a final legal determination of the relevant legal issues.³⁴⁷

According to the Chamber, 'it has now been unequivocally established, both domestically and by [the ICC], that South Africa must arrest Omar Al-Bashir and surrender him to the Court ... any possible ambiguity as to the law concerning South Africa's obligations has been removed ...'.³⁴⁸ It suggests that 'the present decision comprehensively and conclusively disposes of the matter as concerns South Africa's obligations under the Rome Statute'.³⁴⁹

³⁴⁶ Ibid at paras 127–134.

³⁴⁷ Ibid at para 139.

³⁴⁸ Ibid at para 137.

³⁴⁹ Ibid at para 136.

Neither Complimentary nor Complementary: *National Commissioner of the South African Police Service v Southern African Litigation Centre and Another*¹

Salim A Nakhjavani*

The Rome Statute of the International Criminal Court (ICC)² was the hard-won prize of over a century of gestation,³ a few fits and starts⁴ and a final marathon negotiating session in June 1998.⁵ Aside from the powers of its independent Prosecutor, the relationship between the ICC and national justice systems was one of the most complex and fraught aspects of the negotiations. In the final text, States agreed that the ICC should be one of last resort, acting when States themselves are ‘unwilling or unable’⁶ to exercise their primary responsibility to prevent impunity for the most serious (so-called ‘core’) international crimes: genocide, crimes against humanity, war crimes and, after more time and further negotiations,⁷ aggression. The concept of *complementarity*⁸ emerged as the byword

* Adjunct Professor, University of the Witwatersrand; Pupil Member of the Johannesburg Bar.

¹ [2014] ZACC 30, 2015 (1) SA 315 (CC), 2015 (1) SACR 255 (CC), 2014 (12) BCLR 1428 (CC).

² Rome Statute of the International Criminal Court, 2187 UNTS 90, entered into force 1 July 2002 (‘Rome Statute’).

³ See, eg, CK Hall ‘The First Proposal for A Permanent International Criminal Court’ (1998) 322 *International Review of the Red Cross* 57 (Attributing the first proposal for a permanent international criminal court to Gustave Moynier in 1872, in connection with the atrocities of the Franco-Prussian war). For a more extensive treatment, see C Çakmak ‘Evolution of the Idea of a Permanent International Criminal Court Prior to World War I’ (2008) 4 *Review of International Law and Politics* 135.

⁴ See, eg, K Ambos *Treatise on International Criminal Law: Volume 1: Foundations and General Part* (2013) 16–19 (Reviewing the efforts of the International Law Commission, International Law Association and Association International de Droit Pénal in the period 1947–1994).

⁵ For two useful, ‘inside’ perspectives of the Rome Conference, see P Kirsch & JT Holmes ‘The Rome Conference on an International Criminal Court: The Negotiating Process’ (1999) 93(1) *American Journal of International Law* 2; and J Washburn ‘The Negotiation of the Rome Statute for the International Criminal Court and International Lawmaking in the 21st Century’ (1999) 11(2) *Pace International Law Review* 361. See also ‘Official Records of the United Nations Diplomatic Conference of Plenipotentiaries on the Establishment of an International Criminal Court’ (Rome, 15 June–17 July 1998), available at <http://legal.un.org/icc/rome/proceedings/contents.htm>.

⁶ See Rome Statute art 17.

⁷ Assembly of States Parties ‘The Crime of Aggression’ Resolution RC/Res.6 (adopted 11 June 2010) (This amendment was adopted at the Review Conference mandated by Rome Statute art 123 held at Kampala, Uganda in 2010).

⁸ See, eg, Rome Statute preamble, 10th recital: ‘Emphasizing that the International Criminal Court established under this Statute shall be complementary to national criminal jurisdictions.’

for structuring the *vertical* relationship between the ICC and its members – currently numbering 124 States Parties, including (at the time of writing) South Africa,⁹ and significantly, excluding Zimbabwe.

What the architects of complementarity in Rome did not anticipate – not, at least, on a search of the written record – was the extension of the concept of complementarity to regulate *horizontal* relationships between states, including states parties.¹⁰ Neither does the ICC itself: in 2009, its Pre-Trial Chamber defined complementarity in relation to the admissibility of cases before the ICC, and in strictly *vertical* terms:

Complementarity is the principle reconciling the States' persisting duty to exercise jurisdiction over international crimes with the establishment of a permanent international criminal court having competence over the same crimes; admissibility is the criterion which enables the determination, in respect of a given case, whether it is for a national jurisdiction or for the Court to proceed.¹¹

Yet a horizontal application of the Rome Statute's complementarity standard is exactly what the Constitutional Court mandates, by unanimous judgment, in *National Commissioner of the South African Police Service v Southern African Litigation Centre and Another (SALC)*.¹²

This judgment has already been lauded as having 'breathed new life into the principle of universal jurisdiction',¹³ by providing 'clarification of [its] scope' and permitting investigation but not prosecution *in absentia*.¹⁴ The praise is not merely external. From the perspective of discourse analysis, the Court's reasoning itself speaks from a position of not only legal authority but a veritable moral

⁹ On 19 October 2016, South Africa announced its intention to withdraw from the Rome Statute, depositing its Instrument of Withdrawal under art 127(1) of the Rome Statute. The Instrument was rescinded on 7 March 2017, following the decision of the Full Bench of the High Court on 22 February 2017 that *both* the implementation of the decision to withdraw by the Minister of International Relations and Cooperation, and the depositing of the instrument of withdrawal with the Secretary General of the United Nations without prior parliamentary approval were unconstitutional and invalid. *Democratic Alliance v Minister of International Relations and Cooperation and Others (Council for the Advancement of the South African Constitution Intervening)* [2017] ZAGPPHC 53, 2017 (3) SA 212 (GP), [2017] 2 All SA 123 (GP) at para 84. It is unclear whether South Africa will renew the withdrawal process in the future.

¹⁰ See C. Ryngaert 'Complementarity in Universality Cases: Legal-Systemic and Legal Policy Considerations' in M. Bergsmo (ed) *Complementarity and the Exercise of Universal Jurisdiction for Core International Crimes* (2010) 165 ('[T]he complementarity principle, as designed by the drafters of the Rome Statute, was meant to apply vertically. Vertical complementarity means that a supranational institution, the International Criminal Court (ICC), would supervise the investigative and prosecutorial work of States (Parties to the Rome Statute), and, applying Article 17 of the Statute, assume its responsibilities (that is, declare a case admissible) if that work proved to be below acceptable standards.')

¹¹ *Prosecutor v Joseph Kony et al., Decision on the Admissibility of the case under article 19(1) of the Statute* ICC-02/04-01/05-377, Pre-Trial Chamber II (10 March 2009) at para 34.

¹² Note 1 above.

¹³ A. Mudukuti 'The Zimbabwe Torture Case: Reflections on Domestic Litigation for International Crimes in Africa' in S. Williams & H. Woolaver (eds) *Civil Society and International Criminal Justice in Africa: Challenges and Opportunities* (2016) 287, 288, also published in (2016) *Acta Juridica* 287.

¹⁴ H. Woolaver 'Partners in Complementarity: The Role of Civil Society in the Investigation and Prosecution of International Crimes in South Africa' in Williams & Woolaver (note 13 above) 129 at 140–141.

righteousness, not uncommon in international crimes prosecutions¹⁵ and, indeed, from a critical perspective, a characteristic of mainstream international criminal law:¹⁶

Our country's international and domestic law commitments must be honoured. We cannot be seen to be tolerant of impunity for alleged torturers. We must take up our rightful place in the community of nations with its concomitant obligations. We dare not be a safe haven for those who commit crimes against humanity.¹⁷

Passages such as these leave the reader with the impression of a 'hurrah' judgment, reflecting the recent concerns of critical scholarship in the field:

International criminal law has (both formally and rhetorically) been instrumental in the designation of 'outlaw states'. ... Increasingly the enforcement of international criminal law has become the yardstick against which states are measured and sovereign privilege is granted or revoked (see, for one, complementarity). One might suggest that increasingly international crimes are doing the rhetorical work that the notions of human rights and development can no longer undertake as effectively after years of sustained critique.¹⁸

One need not agree with these and similar views to acknowledge that they offer valuable and challenging insights for the exploration of social reality. Gevers, for instance, goes on to suggest – presciently, in light of South Africa's potential withdrawal from the Rome Statute – that the stigmatising 'anti-pluralist' and hegemonic undertow of international criminal law has proved to be the principal locus of struggle for African states parties to the Rome Statute.

The legal standard for assessing complementarity, captured in art 17 of the Rome Statute,¹⁹ is in reality the warp and woof of the entire legal regime – the

¹⁵ F Mégret 'Accountability and Ethics' in L. Reydam, J. Wouters & C. Ryngaert *International Prosecutors* (2012) 417.

¹⁶ C. Schwöbel 'The Market and Marketing Culture of International Criminal Law' in C. Schwöbel (ed) *Critical Perspectives on International Criminal Law* (2014) 270.

¹⁷ *SALC* (note 1 above) at para 80.

¹⁸ C. Gevers 'International Criminal Law and Individualism' in Schwöbel (note 16 above) 228–229.

¹⁹ Rome Statute art 17 reads in full:

'1. Having regard to paragraph 10 of the Preamble and article 1, the Court shall determine that a case is inadmissible where:

- (a) The case is being investigated or prosecuted by a State which has jurisdiction over it, unless the State is unwilling or unable genuinely to carry out the investigation or prosecution;
- (b) The case has been investigated by a State which has jurisdiction over it and the State has decided not to prosecute the person concerned, unless the decision resulted from the unwillingness or inability of the State genuinely to prosecute;
- (c) The person concerned has already been tried for conduct which is the subject of the complaint, and a trial by the Court is not permitted under article 20, paragraph 3;
- (d) The case is not of sufficient gravity to justify further action by the Court.

2. In order to determine unwillingness in a particular case, the Court shall consider, having regard to the principles of due process recognized by international law, whether one or more of the following exist, as applicable:

- (a) The proceedings were or are being undertaken or the national decision was made for the purpose of shielding the person concerned from criminal responsibility for crimes within the jurisdiction of the Court referred to in article 5;
- (b) There has been an unjustified delay in the proceedings which in the circumstances is inconsistent with an intent to bring the person concerned to justice;

‘cornerstone’, to adopt the language of the ICC Pre-Trial Chamber.²⁰ Through a complex internal logic of provisions of the Rome Statute and its Rules of Procedure and Evidence, this *complementarity test* serves as a central reference point in guiding: the exercise of the Prosecutor’s power to initiate investigations;²¹ the admissibility of situations;²² the admissibility of cases;²³ review of determinations of inadmissibility of cases;²⁴ the confirmation of charges before the Pre-Trial Chamber;²⁵ and the myriad other circumstances in which admissibility is either challenged or falls to be determined *mero motu*.

The core of the judicial reasoning in *SALC* is open to critique because it transposes the normative content of the complementarity test (from the Rome Statute via the ICC Act),²⁶ where it is intended to regulate *vertical relationships* between the ICC and states parties, and applies it outside the ICC framework to *horizontal relationships* between states, without the procedural safeguards the Rome Statute provides. These safeguards include, most notably, a prominent role in proceedings for the ‘home states’ (that is, the states of nationality of the person under investigation, and on whose territory the alleged crimes occurred); the uncontested availability of the defence *ne bis in idem*; and the availability of a ‘complementarity arbiter’ acceptable to both the forum and home States, namely the ICC itself, by means of their ratification of the Rome Statute regime.

By contrast, the record in *SALC* shows that the Zimbabwean authorities were not approached by the victims, although with apparently good reason;²⁷ there is no indication that South Africa provided notice of the proceedings to Zimbabwe or afforded that State an opportunity to act,²⁸ even in a manner that would safeguard the confidentiality and security of the complainants; there was no role for Zimbabwe in proceedings in which the Constitutional Court finds the authorities of that country either ‘unwilling’ or ‘unable’ to investigate or

(c) The proceedings were not or are not being conducted independently or impartially, and they were or are being conducted in a manner which, in the circumstances, is inconsistent with an intent to bring the person concerned to justice.

3. In order to determine inability in a particular case, the Court shall consider whether, due to a total or substantial collapse or unavailability of its national judicial system, the State is unable to obtain the accused or the necessary evidence and testimony or otherwise unable to carry out its proceedings.’

²⁰ *Joseph Kony* (note 11 above) at para 34.

²¹ See Rome Statute art 15, read with rule 48, which refers to art 53(1)(b), which refers back to art 17.

²² See Rome Statute art 18(3), which refers to a ‘change of circumstances based on a State’s unwillingness or inability to genuinely carry out the investigation’ – the wording of the complementarity test in art 17.

²³ See Rome Statute art 19(1) which refers back to art 17.

²⁴ See Rome Statute art 19(10) which refers back to art 17.

²⁵ See Rules of Procedure and Evidence rule 122, which refers to rule 58.

²⁶ Implementation of the Rome Statute of the International Criminal Court Act 27 of 2002 (‘ICC Act’).

²⁷ *SALC* (note 1 above) at para 62 (‘Zimbabwe was not asked by the alleged victims of torture to investigate the crime’).

²⁸ See J Stigen ‘The Relationship between the Principle of Complementarity and the Exercise of Universal Jurisdiction for Core International Crimes’ in Bergsmo (note 10 above) 133 at 158 ff (Assessment of proposals that international law allows the forum State to offer the case to the home State or States as a means of ‘proactive subsidiarity’).

prosecute its nationals for acts of torture;²⁹ there is no binding principle *ne bis in idem* in a transnational context, without which the rights of any person eventually prosecuted in South Africa would be in (double) jeopardy;³⁰ and finally, Zimbabwe does not share the Rome Statute framework, or indeed have recourse to any other impartial ‘complementarity arbiter’ in relation to any eventual investigations or prosecutions in South Africa.³¹

The logic of Majiedt AJ, writing for a unanimous bench, is generally straightforward and can be sketched out in brief: from the proposition that torture is an international crime,³² he draws the entirely correct conclusion that South Africa has international customary and treaty obligations to prosecute the crime of torture.³³ He goes on to find that while physical presence of the alleged torturer may be required for a prosecution to proceed, presence on South African territory is not a legal requirement at the investigation stage, being an exercise of *adjudicative* and not *enforcement* jurisdiction.³⁴ He then finds that three cumulative legal requirements must be satisfied to justify the exercise of universal jurisdiction at the investigation stage – in other words, that the exercise of universal jurisdiction by South Africa under customary international law (and its own ICC Act) is limited in three ways:

- a. firstly, a requirement of *subsidiarity*,³⁵ namely that there exists a ‘substantial and true connection between the subject-matter and the source of the jurisdiction’;³⁶
- b. second, a requirement of *complementarity*,³⁷ namely that the territorial (and presumably the national State) are unwilling or unable genuinely to investigate or prosecute; and
- c. third, a case-by-case assessment of the *practicality*³⁸ of the investigation sought.

Admittedly, a degree of confusion is introduced into the reasoning in one problematic paragraph, where considerations of subsidiary and complementarity appear to be conflated. The reasoning here seems to construe complementarity either as a subset of subsidiarity, or as a means of limiting jurisdiction previously founded on the basis of subsidiarity.³⁹ On this basis, Majiedt AJ refers to ‘at least two limitations’: subsidiarity and practicality.⁴⁰ A close reading of the later paragraphs, however, suggests that the Court considered complementarity as a third,

²⁹ *SALC* (note 1 above) at para 62.

³⁰ Ryngaert (note 10 above) at 170–172. See also C van Den Wyngaert & G Stessens ‘The International Non Bis In Idem Principle: Resolving Some of the Unanswered Questions’ (1999) 48 *International and Comparative Law Quarterly* 779–804.

³¹ The precarious neutrality of the forum State in assessing complementarity standards in the home State is explored in some depth by Stigen (note 28 above) at 157.

³² *SALC* (note 1 above) at para 38.

³³ *Ibid* at para 40.

³⁴ *Ibid* at para 29.

³⁵ *Ibid* at paras 61–62, 78.

³⁶ *Ibid* at para 61. Compare *ibid* at para 28, which refers to a ‘substantial and bona fide connection’.

³⁷ *Ibid* at paras 61–62 and 78.

³⁸ *Ibid* at para 63.

³⁹ *Ibid* at para 61.

⁴⁰ *Ibid*.

self-standing limitation on the exercise of universal jurisdiction: while it frames the subsidiarity threshold ('substantial and true connection') as establishing a jurisdictional nexus between South Africa and alleged crimes committed abroad, it characterises the complementarity test ('unwilling or unable to prosecute') as an expression of the principle of non-intervention.⁴¹ These two are expressions of related but distinct principles of international law. Subsidiarity works to found jurisdiction and prevent jurisdictional overreach; complementarity respects the UN Charter-based principle of non-intervention in the internal affairs of other States.

The idea that subsidiarity 'founds' jurisdiction is an intriguing and valuable one that receives inadequate attention in the judgment, and would have strengthened its reasoning. This is because it dispels a longstanding misunderstanding in the literature and in practice about the nature of universal jurisdiction, which does not itself *establish* or *found* jurisdiction but merely *describes* a set of circumstances, framed as a negative or residual category,⁴² under which States, as a matter of international law, are *permitted but not required* to exercise criminal jurisdiction over core international crimes. In other words, permissive universal jurisdiction is 'jurisdiction exercised over crimes committed abroad where there are no links of nationality to the suspect or victim or of harm to the state's own special interests'.⁴³

But states choosing to exercise permissive universal jurisdiction remain bound by other norms of international law in taking action: the principle of non-intervention, for instance. Universal jurisdiction does not somehow suspend the operation of the international legal system; it is part of that complex, adaptive system, in which it plays a part that occasionally defies linear prediction.

To regulate the inevitable tensions that arise between those rules of international law that tend to entrench State sovereignty and those that promote international justice, a number of legal balancing tests have evolved in the practice of states. Majiedt AJ rightly identifies three: complementarity, subsidiarity and practicality. Without stretching the metaphor, these tests play the role of potentiometers in the international circuitry: they regulate the flow of (state) power by varying (judicial) resistance.

However, the judgment itself works at cross-purposes on this point. While founding jurisdiction on the principle of subsidiarity in one place,⁴⁴ an earlier passage misconstrues universal jurisdiction – whether intentionally or through lax usage – by appearing to make its exercise dependent – in practice if not in law – on membership of the Rome Statute regime:

⁴¹ Ibid.

⁴² R O'Keefe 'Universal Jurisdiction: Clarifying the Basic Concept' (2004) 2 *Journal of International Criminal Justice* 735, 745 and fn 12.

⁴³ CK Hall 'The Role of Universal Jurisdiction in the International Criminal Court Complementarity System' in Bergsmo (note 10 above) 201 at 205. See also *ibid* at 202: ('[U]niversal jurisdiction means the *ability* of the court of any state to try persons for crimes committed outside its territory which are not linked to the forum state by the nationality of the suspect or of the victims at the time of the crimes or by harm to that state's own special national interests.')

⁴⁴ *SALC* (note 1 above) at para 61.

If an investigation is not instituted by non-signatory countries in which the crimes have been committed, the perpetrators can only be brought to justice through the application of universal jurisdiction, namely the investigation and prosecution of these alleged crimes *by states parties under the Rome Statute*.⁴⁵

This is a category mistake. The exercise of universal jurisdiction is conceptually distinct from South Africa's Rome Statute obligations. Indeed, many states prescribe the exercise of universal jurisdiction over a range of so-called 'ordinary' (non-Rome Statute) crimes.⁴⁶ The Court's reliance on the ICC Act also works at cross-purposes. While the Court grounds the investigative powers of SAPS over the alleged instances of torture in this instance on the ICC Act,⁴⁷ the only valid basis on which South Africa and Zimbabwe share reciprocal obligations under international law to investigate and prosecute torture is as a crime under peremptory norms of customary international law, as well as under the Torture Convention. The ICC Act enacts Rome Statute crimes into South African law, including torture *only when committed as a crime against humanity* – that is, in the context of a widespread or systematic attack directed against a civilian population. It does not, by any means, domesticate the Torture Convention or the customary international law prohibition on torture (a measure effected instead by the Torture Act).⁴⁸ The judgment's own analysis of the purposes of the ICC Act makes no such sweeping finding.⁴⁹ The character of subsidiarity as founding jurisdiction does not receive sufficient attention in the judgment itself.

It is equally noteworthy that the language of the subsidiarity threshold – 'substantial and true connection' – adopted by the Court here echoes the earlier pronouncement of Sachs J in *S v Basson*, also in the context of core international crimes:

'to make an offence subject to the jurisdiction of our courts ... it is sufficient that there be a "real and substantial link" between an offence and this country, a test well-known in public and private international law.'⁵⁰

Nonetheless, the judgment's only tangential reference to *Basson*⁵¹ does not include any consideration of this central pronouncement on the principle of subsidiarity in its application to core international crimes. With this omission, the Court loses an opportunity not only to further develop its own jurisprudence from *Basson* but to harmonise the muddled state of public international law on subsidiarity in the context of universal jurisdiction with its analogue from private international law, which is underpinned by a depth of relatively stable and consistent comparative jurisprudence.

⁴⁵ Ibid at para 32 (emphasis added).

⁴⁶ See Hall (note 43 above) at 205–206 (Providing examples of the exercise of universal jurisdiction over ordinary crimes such as murder, rape, assault or abduction).

⁴⁷ See *SALC* (note 1 above) at paras 54–60.

⁴⁸ Prevention of Combating and Torture of Persons Act 13 of 2013, read with s 231(4) of the Constitution.

⁴⁹ *SALC* (note 1 above) at paras 33–34.

⁵⁰ *S v Basson* [2005] ZACC 10, 2007 (3) SA 582 (CC), 2005 (12) BCLR 1192 (CC) at para 226, quoting *Libman v The Queen* [1985] 2 SCR 178 at 212–213 (LaForest J).

⁵¹ *SALC* (note 1 above) at para 30, fn 24 (Citing with approval the reasoning of Sachs J concerning the ongoing responsibility of states to try cases of breaches of international humanitarian law).

With all due deference, the judgment is strikingly under-researched and thus insufficiently reasoned in at least four additional ways, even allowing for the complexity of the case, the reality of judicial time-pressures in our apex court,⁵² and the resources of four senior and 11 junior counsel for the parties.

Firstly, although the judgment quotes the Joint Separate Opinion of Judges Higgins, Kooijmans and Buergenthal in the *Arrest Warrant* case⁵³ before the International Court of Justice in support of the proposition that physical presence on the territory of the forum state is not a precondition for an investigation,⁵⁴ it neither cites nor judicially considers the *most directly relevant legal finding* of that same Opinion: ‘A State contemplating bringing criminal charges based on universal jurisdiction must first offer to the national State of the prospective accused person the opportunity itself to act upon the charges concerned.’⁵⁵

Second, the judgment makes no reference, either to approve or disapprove the line of ICC decisions from 2009 and 2010 that apply the complementarity test in diverse contexts and, in particular, begin to elucidate the legal approach and factors to be considered in assessing unwillingness or inability under the Rome Statute framework.⁵⁶

Third, to the extent consideration of state practice as an element of customary international law is mandated by s 233 of the Constitution, the judgment makes no mention of the limited but instructive foreign decisions relating to subsidiarity in the context of the exercise of universal jurisdiction. These decisions include: the Spanish Supreme Court’s application of stringent subsidiarity rule in the 2003 *Guatemalan Genocide* case;⁵⁷ the Spanish Constitutional Court’s 2005 ruling that the exercise of universal jurisdiction is limited only by a *ne bis in idem* safeguard, not any rule of subsidiarity;⁵⁸ the Spanish Appeal Court’s 2009 reversal of a decision to prosecute in the *Al-Daraj* case, finding it ‘inadmissible to question the competence of the judicial authorities of the State of Israel to investigate, and if fitting, to try the events’;⁵⁹ and the published decisions of German prosecution authorities from 2005 and 2007 not to proceed with investigations in the *Abu Ghraib* prison abuse matter, first on subsidiarity alone and subsequently on an amalgam of subsidiarity and practicality of the investigation, finding that: ‘The view of the complainant that the Federal Republic of Germany must act as a

⁵² A period of slightly over five months elapsed between hearing (19 May 2014) and judgment (30 October 2014).

⁵³ Case Concerning the Arrest Warrant of 11 April 2000 (*Democratic Republic of the Congo v Belgium*) [2002] ICJ 3 (*Arrest Warrant*) (14 February 2002).

⁵⁴ *SALC* (note 1 above) at fn 50.

⁵⁵ *Arrest Warrant* (note 53 above) at para 59.

⁵⁶ Among these, *Prosecutor v Germain Katanga and Mathieu Ngudjolo Chui, Judgment on the Appeal of Mr. Germain Katanga against the Oral Decision of Trial Chamber II of 12 June 2009 on the Admissibility of the Case* ICC-01/04-01/07-1497 (25 September 2009) at paras 78 and 85; Joseph Kony (note 11 above) at paras 45 and 51; *Situation in the Republic of Kenya, Decision Pursuant to Article 15 of the Rome Statute on the Authorization of an Investigation into the Situation in the Republic of Kenya* Case No. ICC-01/09-19 (31 March 2010) at paras 50 and 183; *The Prosecutor v Jean-Pierre Bemba Gombo, Decision on the Admissibility and Abuse of Process Challenges* (*Bemba Admissibility Decision*) Case no. ICC-01/05-01/08-802 (24 June 2010) at paras 239–247.

⁵⁷ Tribunal Supremo Case No. 327-2003 (25 February 2003).

⁵⁸ Tribunal Constitucional Case No. STC 237/2005 (26 September 2005).

⁵⁹ Audiencia Nacional Appeal No. 31/09 (9 July 2009).

representative of the “international community” and therefore at least take up investigations is thus mistaken.⁶⁰

Fourth, the judgment would have been strengthened by explicit reference or at least some incorporation of argument from the work of Rastan, Stigen or Ryngaert – each characterised in 2010 as ‘leading international experts’⁶¹ on complementarity, and each having analysed the specific issue of the application of the subsidiarity principle in the exercise of universal jurisdiction in domestic legal systems.

Both individually and cumulatively, these are deafening silences.

How, then, should the Court have approached subsidiarity and complementarity in *SALC*? Scholars and practitioners alike recommend both caution and depth of reasoning. More prosaically: ‘If you do it, do it right.’ Cedric Ryngaert makes the following targeted observation:

[T]here are strong normative arguments in favour of a principle of horizontal complementarity, although admittedly it may not yet have crystallized as a norm of customary international law given the dearth of pertinent state practice. However, stating that there is such a thing as horizontal complementarity is one thing, implementing it correctly is quite another. A warning may have to be provided here as to an overly policy-based horizontal complementarity analysis. Lacking principled guidance, such an analysis may easily be contorted for political purposes. And because prosecutors are not under a legal duty to carry out a complementarity analysis, assuming that there are no administrative guidelines on horizontal complementarity which are binding on them either, they may even believe that they can do wholly without a complementarity analysis, or at least carry out a very superficial self-serving analysis without genuinely inquiring into whether the territorial or national state has conducted any relevant proceedings.⁶²

It is worth recalling, in this vein, that the Court’s ‘unwillingness or inability’ analysis of the Zimbabwean judicial system is limited to a brief paragraph, and that the home state was neither notified of the proceedings nor invited to have its views heard.

As Jo Stigen has observed, compellingly:

There is, however, an inherent paradox with the application of such a subsidiarity criterion. Absent an international scrutiny mechanism, it presupposes a horizontal scrutiny between states of the adequacy of their respective proceedings. This is quite different from the vertical scrutiny exercised by the ICC. Thus, while initially aiming at reducing the risk of interstate friction, subsidiarity can also make the application of universal jurisdiction more intrusive. This makes it all the more important that the most essential aspects of the complementarity principle aimed at safeguarding the integrity of states vis-à-vis the ICC are applied *mutatis mutandis* to the exercise of universal jurisdiction.⁶³

Chief among these safeguards is the notification of the home state by the forum state and the offer of the case for genuine investigation, with support if necessary,

⁶⁰ Decision 3 ARP 156/06-2 (5 April 2007). See also Decision 3 ARP 207/04-2 (10 February 2005).

⁶¹ M Bergsmo ‘Between Territoriality and Universality: Room for Further Reflection’ in Bergsmo (note 10 above) 1.

⁶² Ryngaert (note 10 above) at 190.

⁶³ Stigen (note 31 above) at 158.

in a manner that guarantees the confidentiality and safety of complainants.⁶⁴ The guarantee of *ne bis in idem* protection as between the forum and home states would also be a requirement of international human rights law, at least once the investigation crystallises to the point that suspects are 'substantially affected' by the suspicion against them, or formally notified that they are suspects in the investigation.⁶⁵

It may be helpful, in conclusion, to make explicit that the efflorescence of the principle of horizontal complementarity is indicative of a broader shift in the discourse of international criminal law. The language of a 'web' of universal jurisdiction as a foil to the insularity of a corrupted national sovereignty, ensnaring perpetrators of international crimes, was prevalent around the adoption and entry into force of the Rome Statute – and is still apparent today in the language of 'no safe haven'.⁶⁶ But the discourse has matured and reflects a deeper understanding of the challenges of *managing overlapping responsibilities*: the *primary responsibility* of the home states (typically on the basis of territoriality or nationality of the offender) and the *secondary responsibility* of other states as well as international courts and tribunals. Quite beautifully, this shift both reinforces and recasts the value of sovereignty in regulating world order.

In light of these comments, it may also be helpful to conclude with an examination of the means by which the judgment seeks to legitimate the exercise of universal jurisdiction. What is the Constitutional Court saying about South Africa's (and indeed, Zimbabwe's) responsibilities? Whether one understands complementarity only in its narrow, vertical sense, or admits a broader notion of complementarity as 'burden-sharing' in the fight against impunity for atrocity crimes, not only between international courts and domestic legal systems but also between states,⁶⁷ the Court's legal characterisation of the facts under the subsidiarity test is reduced to one sentence of reasoning: 'Given the international and heinous nature of the crime, South Africa has a substantial connection to it.'⁶⁸

But to displace the jurisdictional claim of the home states – the states of territoriality and nationality, which have traditionally been accorded some priority out of pragmatism if not in binding international law⁶⁹ – requires legitimisation. In the context of horizontal complementarity for core international crimes, legitimisation must arise from a threshold – subsidiarity or otherwise – that *actually means something*. To say that every core international crime automatically bears a 'substantial and true' connection to South Africa, as the judgment does, is to render the subsidiarity threshold nugatory and thus irrelevant.

I am indebted to the anonymous reviewer who took the point that universal jurisdiction is unashamedly normative in character. A thorough exploration of the

⁶⁴ See the detailed analysis of (note 31 above) at 151–153.

⁶⁵ *Deveer v Belgium* [1980] ECHR 1 at paras 42, 44 and 46.

⁶⁶ See, eg, *SALC* (note 1 above) at para 81.

⁶⁷ R Rastan 'Complementarity: Contest or Collaboration?' in Bergsmo (note 10 above) 83, 83–84 and 106 ff.

⁶⁸ *SALC* (note 1 above) at para 78.

⁶⁹ Rastan (note 67 above) at 99.

theory of representation that legitimates claims to universal jurisdiction lies beyond the scope of this comment.

By way of a concluding *excursus*, however, the recent pronouncement of the full bench of the High Court, declaring South Africa's attempted withdrawal from the Rome Statute unconstitutional and invalid, does contain a curious, thought-provoking and challenging *obiter dictum* that warrants our collective attention:

Therefore, the approval of an international agreement in terms of s 231(2) creates a social contract between the people of South Africa, through their elected representatives in the legislature, and the national executive. That social contract gives rise to the rights and obligations expressed in such international agreement.⁷⁰

If we accept that 'social contract' in this context need not refer solely to a state-bound or nationally constructed society, but to all the conceptual and lived richness that characterises Allott's theory of human self-constituting,⁷¹ this seemingly offhand remark from the High Court takes on significant meaning. It may be that international law lifts the 'state veil' to some extent. It may be that the legislative, representative function, rather than the executive one, legitimates the exercise of jurisdiction under international law. In a prescient article, Hume argued that the fully fledged 'real and substantial link' test, transposed from private to public international law, is 'a *constitutive* element of *Parliament's legislative competence under public international law*...[it] legitimates the exercise of *Parliament's authority on the international level*'.⁷² This argument is especially compelling because it identifies with precision the nature of the juridical link being created between the forum state exercising universal jurisdiction, and the alleged criminal conduct: it is the *legislature* extending its will to the international plane – not as a mere agent of the international community, but in its own right.

The matter of South Africa's attempted withdrawal from the Rome Statute will not be appealed by the state. Further judicial engagement on the vital question of horizontal complementarity will have to await a future test case.

⁷⁰ *Democratic Alliance* (note 9 above) at para 52.

⁷¹ P Allott 'Globalization from Above: Actualizing the Ideal through Law' in K Booth, T Dunne & M Cox (eds) *How Might we Live? Global Ethics for a New Century* (2001) 61ff. See also P Allott *Eunomia: New Order for a New World* (1990).

⁷² N Hume 'Four Flaws: Reflections on the Canadian Approach to Private International Law' (2006) 44 *Canadian Yearbook of International Law* 161, 217.

Articles

Constitutional Heedlessness and Over-Excitement in the Common Law of Delict's Development

Emile Zitzke*

I INTRODUCTION

In this piece I take issue with two seemingly contradictory ways of approaching the South African common law. On the one hand I problematise the trend of 'constitutional avoidance' (the specific brand of constitutional avoidance that I address here will be called 'constitutional heedlessness') reflected in recent case law relating to the development or application of the common law of delict.¹ On the other hand I also caution against, what I will call, 'constitutional over-excitement'

* Postdoctoral Research Fellow, South African Institute for Advanced Constitutional, Public, Human Rights and International Law (A Centre of the University of Johannesburg). This article is derived from various parts of my doctoral thesis entitled *A New Proposed Constitutional Methodology for Effecting Transformation in the South African Law of Delict* (2016) prepared under the excellent supervision of Professor TJ Scott at the University of Pretoria. My appreciation to Khuraisha Patel, Duard Kleyn, Andrea Bauling and Jason Gouveia, as well as the participants of the Seventh Constitutional Court Review conference for their thought-provoking comments and questions. This article is richer and more coherent thanks to the input of Ngwako Raboshakga and Stu Woolman. Errors remain my own. The financial assistance of the National Research Foundation (NRF) towards this research is hereby acknowledged. Opinions expressed and conclusions arrived at, are those of the author and are not necessarily to be attributed to the NRF.

¹ I acknowledge that 'avoidance' has received much attention in South African scholarship. See, eg, the famous piece by I Currie 'Judicious Avoidance' (1999) 15 *South African Journal on Human Rights* 138 (Critiques the Constitutional Court's decisional minimalism). In this article, I do not delve into the history of avoidance as a broad principle. Instead, I direct my attention to a specific manifestation of avoidance, as it applies to common-law cases, that I call 'constitutional heedlessness', for reasons that I hope will become clear in the course of developing my argument. Constitutional heedlessness is not an unnecessary neologism or synonym for avoidance. There are other manifestations of avoidance too: there is also the more aggressive form of constitutional avoidance that I have called 'anti-constitutionalism' (see E Zitzke 'A Case of Anti-Constitutional Common-Law Development' (2015) 48 *De Jure* 467), and there is a softer version of constitutional avoidance that I have called a 'constitutionally wanting' approach to the common law's development (see E Zitzke 'Realist Evolutionary Functionalism and Extra-Constitutional Grounds for Developing the Common Law of Delict: A Critical Analysis of *Heroldt v Wills* 2013 2 SA 530 GJ' (2016) 69 *Journal of Contemporary Roman-Dutch Law* 103). Drawing the links between these different brands of avoidance is beyond the scope of this article that is focused on the 2014 term of the Constitutional Court. I leave that for a future endeavor.

which is the polar opposite approach to constitutional heedlessness.² To be clear from the start, by constitutional heedlessness I mean a substantive avoidance of the potential impact of the Constitution of the Republic of South Africa, 1996 on common-law matters that require constitutional infusion, while the courts or authors that employ this approach do not expressly reject the Constitution's potential impact on the matter. In essence, constitutional heedlessness is a 'business-as-usual' approach to the common law – a silent circumvention of the Constitution.³ By constitutional over-excitement I mean a relegation of established common-law rules that are ultimately replaced by a pure application of constitutional principles.

In Part II, I unpack the problem of constitutional heedlessness. Firstly, I explain why constitutional heedlessness is an undesirable paradigm for common-law enquiries. Secondly, I discuss the decisions in *iMvula Quality Protection (Pty) Ltd v Loureiro and Others*,⁴ *H v Kingsbury Foetal Assessment Centre (Pty) Ltd*,⁵ and *Country Cloud Trading CC v MEC, Department of Infrastructure Development, Gauteng*,⁶ as recent manifestations of constitutional heedlessness. *Loureiro SCA* and *Foetal Assessment Centre HC* serve as examples where constitutional heedlessness was remedied in the respective appeals to the Constitutional Court,⁷ while *Country Cloud SCA* serves as an example where, for the purposes of conceptual clarity, a more complete analysis of the relevant issues would have involved a recognition of the constitutional value of state accountability (and the common-law rules that have been developed in light of this norm), even though such recognition did not have an effective impact on Country Cloud's appeal to the Constitutional Court.⁸ In Part III, I turn to the competing approach of constitutional over-excitement portrayed in the way in which the wrongfulness enquiry was addressed in *Loureiro CC*. Finally, in Part IV, I develop the argument that even though it may be desirable to take the Constitution seriously so that it militates against common-law veneration and its ideological stagnation, the Constitution should not be monumentalised to such an extent that we become uncritical of it. I further contend that the doctrine of adjudicative subsidiarity may provide useful conceptual machinery to strike a balance between the two extreme approaches at issue in this discussion.

² Even though other scholars have critiqued courts for wrecking the common law in favour of constitutional principles in the past, I use the term 'constitutional over-excitement' as one that aims to unify recurrent critiques of this nature. I by no means suggest that I am the first person to criticise extreme zealotry in constitutional application to the common law. I mention relevant scholarship that has done similar work in delict in Part III below. For the purposes of developing a juxtaposed critique of two opposing problems, I find the new term (for an old problem) to be both useful and necessary.

³ D Davis *Democracy and Deliberation: Transformation and the South African Legal Order* (1999) 127–163 (Problematises the lack of constitutional imagination in private-law scholarship).

⁴ [2013] ZASCA 12, 2013 (3) SA 407 (SCA), [2013] 2 All SA 659 (SCA) ('*Loureiro SCA*').

⁵ [2014] ZAWCHC 61 ('*Foetal Assessment Centre HC*').

⁶ [2013] ZASCA 161, 2014 (2) SA 214 (SCA) ('*Country Cloud SCA*').

⁷ *Loureiro and Others v iMvula Quality Protection (Pty) Ltd* [2014] ZACC 4, 2014 (3) SA 394 (CC), 2014 (5) BCLR 511 (CC) ('*Loureiro CC*'); *H v Fetal Assessment Centre* [2014] ZACC 34, 2015 (2) SA 193 (CC), 2015 (2) BCLR 127 (CC) ('*Fetal Assessment Centre CC*').

⁸ *Country Cloud Trading CC v MEC, Department of Infrastructure Development, Gauteng* [2014] ZACC 28, 2015 (1) SA 1 (CC), 2014 (12) BCLR 1397 (CC) ('*Country Cloud CC*').

II CONSTITUTIONAL HEEDLESSNESS

A The Theoretical Problem

In the middle of Pretoria's Church Square stands a monument dedicated to the late President Paul Kruger, protected by four bronze guards. It is a monument dedicated to a man who once warned that the jurisprudential tradition of natural law was conceived in the womb of the devil.⁹ For Kruger and others, natural law broadly involves the recognition of a higher set of norms against which all laws can be tested. More specifically, for Kruger, the recognition of natural law involved a compromise of the supremacy of the legislature (then called the 'Volksraad') by affording judges a right to test legislation against a higher set of norms.¹⁰ By rejecting the notion of natural law, Kruger aimed to protect the pride of the Volksraad. Under the influence of Kruger's support for parliamentary sovereignty, coupled with the British influence of legal positivism in South African legal scholarship and practice,¹¹ as well as the maintenance of white supremacy and racialised capitalism,¹² the country was in a position to legalise the atrocity of apartheid where a higher set of norms protecting the rights to (among others) life, freedom and security of the person, equality and dignity were unknown to the majority of South Africans subject to oppressive legislation.¹³ If one accepts that the aforementioned rights are all relevant to the natural law tradition,¹⁴ apartheid law involved a clear disapproval of natural law.

At this point it is important to stress that Kruger's stance on natural law specifically related to its application to legislation. Despite the rejection of a natural law theory for statutory interpretation, it appears that many scholars have historically been (and still are) of the view that the rights relevant to the modern developments in the natural law tradition are implicit in the rules of Roman and Roman-Dutch law that form the basis of South African common law.¹⁵ Therefore, many would have regarded (and possibly still would regard) it difficult to attempt to eliminate the natural law tradition from the common law because the latter is inherently pervaded by principles of the former.

Since the eras of Kruger and apartheid a lot has changed in South African law. South Africa now has a supreme Constitution with a justiciable Bill of Rights that

⁹ J Dugard 'Judicial Process, Positivism and Civil Liberty' (1971) 88 *South African Law Journal* 181, 184 (Shows how parliamentary sovereignty and legal positivism upheld apartheid in South Africa and suggests a realist-*cum*-natural law approach instead).

¹⁰ Ibid.

¹¹ Ibid at 184–185.

¹² S Terreblanche *A History of Inequality in South Africa: 1652-2002* (2002) 3–22.

¹³ For an overview, also see ibid at 297ff.

¹⁴ Dugard (note 9 above) at 197.

¹⁵ Ibid at 183. The influential Dutch jurist Hugo de Groot is regarded as the father of Dutch humanism and natural law, which influenced the Roman-Dutch law that was in turn imposed on South Africa. See in this regard DH van Zyl *Geskiedenis van die Romeins-Hollandse Reg* (1979) 191–193; PhJ Thomas, CG van der Merwe & BC Stoop *Historical Foundations of South African Private Law* (2nd Edition, 2000) 57, 70, 72; and WB le Roux 'Natural Law Theories' in C Roederer & D Moellendorf *Jurisprudence* (2004) 40–41. The most recent work on the similarities between the South African common law and human rights is by G van Niekerk 'The Endurance of the Roman Tradition in South African Law' (2011) 4 *Studia Universitatis Babeş Bolyai Jurisprudentia* 20, 21ff (Demonstrates the compatibility of Roman legal principles and constitutional law).

has opened the door for the natural law tradition to thrive in South Africa on all fronts.¹⁶ Furthermore, it is widely accepted today that the Constitution may affect the development of the common law.¹⁷ However, the establishment of a supreme Constitution with a Bill of Rights and its potential impact on the common law was not unequivocally supported by private-law scholars from the start. At the time of democratic transition in South Africa there were some members of the legal academy (and interestingly for present purposes, delict scholars in particular) who took a clear stance against the introduction of a bill of rights or, as a minimum, a stance against the potential infiltration of constitutional rights into the esteemed common law.¹⁸

The rejection of constitutional rights in this context ultimately involved an implicit rejection of a specific brand of the natural law tradition.¹⁹ This is true because it is widely accepted that the institution of human rights is derived from modern developments in natural law theory.²⁰ Therefore, even though President

¹⁶ The formative document that solidified the democratic transition in South Africa, and the concomitant democratic legal reforms, is the Constitution of the Republic of South Africa Act 200 of 1993 ('Interim Constitution'). The Interim Constitution has been repealed.

¹⁷ Since the Constitutional Court's decision in *Du Plessis and Others v De Klerk and Another* [1996] ZACC 10, 1996 (3) SA 850 (CC), 1996 (5) BCLR 658 (CC) ('*Du Plessis*') it has been South African law that the Constitution may have a 'radiating' effect on common law as s 35(3) of the Interim Constitution required that the spirit, purport and objects of the Bill of Rights had to be considered when applying or developing the common law. In *Carmichele v Minister of Safety and Security and Another* [2001] ZACC 22, 2001 (4) SA 938 (CC), 2001 (10) BCLR 995 (CC) ('*Carmichele*') the Court confirmed that common-law developments have to be congruent with the spirit, purport and objects of the Bill of Rights in terms of s 39(2) of the Constitution.

¹⁸ See JM Potgieter 'The Role of the Law in a Period of Political Transition: The Need for Objectivity' (1991) 54 *Journal of Contemporary Roman-Dutch Law* 800 (Raises concerns about the political nature of human rights); PJ Visser & JM Potgieter 'Some Critical Comments on South Africa's Bill of Fundamental Rights' (1994) 57 *Journal of Contemporary Roman-Dutch Law* 493 (Worries about the subjective nature of constitutional adjudication); J Neethling & JM Potgieter 'Laster: Die Bewyslas, Mediaprivilegie en die Invloed van die Nuwe Grondwet' (1994) 57 *Journal of Contemporary Roman-Dutch Law* 513 (Argues that the common law will and should remain largely unaffected by fundamental rights); and the critique of these views by H Botha 'Privatism, Authoritarianism and the Constitution: The Case of Neethling and Potgieter' (1995) 58 *Journal of Contemporary Roman-Dutch Law* 496 (Shows why fundamental rights should have an impact on the common law). See also PJ Visser 'A Successful Constitutional Invasion of Private Law' (1995) 58 *Journal of Contemporary Roman-Dutch Law* 745 (Argues that constitutional interpretation is arbitrary and the infiltration of fundamental rights into the 'empirical' common law will lead to chaos) and the denunciation of those views by G Carpenter & CJ Botha 'The "Constitutional Attack on Private Law": Are the Fears Well Founded?' (1996) 59 *Journal of Contemporary Roman-Dutch Law* 126 (Demonstrates that constitutional interpretation does not have to be arbitrary, vague and embarrassing like Visser alleges).

¹⁹ This is the implication of the critique levelled by Botha, *ibid* at 498.

²⁰ In South African literature, Van Zyl (note 15 above) at 194–195 links the work of De Groot with the thought of Hobbes and Locke (the latter being responsible for the conceptualisation of natural rights). Thomas, Van der Merwe & Stoop (note 15 above) at 111–115 as well as Le Roux (note 15 above) at 41–47 regard these modern developments of natural law theory as instrumental to the establishment of human rights. The link between the natural law tradition and the birth of human rights is largely recognised in works relating to the theory of human rights. See, for example, R McInery 'Natural Law and Human Rights' (1991) 36 *American Journal of Jurisprudence* 1 (Discusses the possibility and limits of a marriage of natural law and human rights); M Discher 'A New Natural Law Theory as a Ground for Human Rights?' (1999–2000) 9 *Kansas Journal of Law & Public Policy* 267 (Reflects on various justifications for the universality of human rights, one of those being its roots in natural law); C Perello 'On Supernatural Law: About the Origins of Human Rights and Natural Law in Antiquity' (2014) 20 *Fundamina* 15 (Links early natural law theory with the concept of human rights); and

Kruger had been dead for roughly 90 years at that stage, his cenotaphic warning against natural law was still being followed, albeit in slightly modified terms. Now natural law, in its human rights form, was to be rejected when it came to working with the South African common law.

The rejection of the infusion of human rights in South African common law is theoretically peculiar. As I have already shown above, the divorce of natural law and common law seems to be a difficult task if it is accepted that natural law is inextricably linked to the rules of common law. As a counter-argument, Hans Visser once favoured such a divorce, contending that the natural law found in Roman-Dutch law is distinct from and superior to the 'backward' and 'savage' hogwash of the 'vague and ambiguous' human rights intended for the South African democratic transformation.²¹ However, Visser eventually relaxed his concerns after he realised that he and similar thinkers had lost the battle against the introduction of fundamental rights in South Africa.²² Visser's colleagues, Johann Neethling and Johan Potgieter, are now proponents of the school of thought that regards common law and human rights as reconcilable, probably because of the shared theoretical foundation of the two sets of rights.²³ If the Roman-Dutch conception of natural law (that underlies the common law) has theoretically developed into human rights and common law continues to be developed through the influence of natural law, then common law and human rights are not only reconcilable but it is also desirable to update the common law in light of human rights.

Beguilingly, the trepidation in delict scholarship regarding the merger of human rights and common law may have slowed down after the Constitution became operative but the anxiety did not come to a complete stop. Even after the widely celebrated pronouncement in *Carmichele* that set the blueprint for a constitutional infusion of common law,²⁴ there have been some delict commentators whose

A Sangiuliano 'Towards a Natural Law Foundationalist Theory of Universal Human Rights' (2014) 5 *Transnational Legal Theory* 218 (Presents a case of 'natural law foundationalism' as the justification for the universality of human rights).

²¹ Visser (note 18 above) at 748–749.

²² See PJ Visser 'Geen Afsonderlike Eis om "Grondwetlike Skadevergoeding" Nie' (1996) 59 *Journal of Contemporary Roman-Dutch Law* 695 (Emphasises that the common law affords a sufficient degree of protection to fundamental rights); and PJ Visser 'Enkele Gedagtes oor die Moontlike Invloed van die Fundamentele Reg op Lewe in die Deliktereg' (1997) 30 *De Jure* 135 (Reflects on various possibilities for positive developments to the common law in light of the Constitution). This turn in Visser's thought is discussed more fully by A van der Walt 'Transformative Constitutionalism and the Development of South African Property Law (Part 1)' (2005) 4 *Journal of South African Law* 655, 661 (Compares the German and South African constitutional approaches to horizontality, specifically in the context of property law).

²³ J Neethling & JM Potgieter *Neethling-Potgieter-Visser Law of Delict* (7th Edition, 2015) 18.

²⁴ *Carmichele* (note 17 above). This decision is celebrated by N Botha 'The Role of International Law in the Development of South African Common Law' (2001) 26 *South African Yearbook of International Law* 253, 259; J Neethling & JM Potgieter 'Toepassing van die Grondwet op die Deliktereg' (2002) 65 *Journal of Contemporary Roman-Dutch Law* 265, 272; M Pieterse 'The Right to be Free from Public or Private Violence after *Carmichele*' (2002) 119 *South African Law Journal* 27, 39; JR Midgley & B Leinius 'The Impact of the Constitution on the Law of Delict' (2002) 119 *South African Law Journal* 17, 27; G Carpenter 'The Carmichele Legacy – Enhanced Curial Protection of the Right to Physical Safety' (2003) 18 *South African Public Law* 252, 266; J Neethling 'Die Carmichel-Sage Kom tot 'n Gelukkige Einde' (2005) 2 *Journal of South African Law* 402, 409; and D Davis & K Klare 'Transformative

work show moments of discomfort with the way in which the Constitution has been used to transform the common law. These commentators do not appear to be writing from a position of political panic about the democratisation of South Africa as Visser once was. Rather, they are concerned with the legal technicalities of whether the Constitution could and should add substance to the common law.

The clearest stance against a constitutional colonisation of the common law is found in the latest work by Johan van der Walt. Van der Walt, who once seemed enthusiastic about the potential of *Carmichele*,²⁵ later expressed the view that the common law could have been able to provide Ms Carmichele with the necessary relief against the state's negligence because the common law recognised the assortment of rights relating to bodily integrity.²⁶ Most recently, Van der Walt has taken a radical turn by rejecting the infiltration of constitutional reasoning in common-law matters except for certain exceptional circumstances where a counter-majoritarian difficulty arises.²⁷ Concisely, it is Van der Walt's stance that the common law can provide enough protection to the rights of parties without necessarily invoking the Constitution.²⁸ Even though Van der Walt's position is closer to an anti-constitutional strategy for the common law, he would certainly not be opposed to courts employing a constitutionally heedless approach when dealing with common law.

A less radical stance of constitutional reservation is reflected in Anton Fagan's philosophy of common-law development. Even though Fagan does not appear to be completely opposed to the essential idea of constitutional scrutiny of the common law,²⁹ he contends that both *Carmichele*³⁰ and *K v Minister of Safety and Security*³¹ were incorrectly decided as far as the interaction between the common law and the Constitution is concerned. Drawing from a joint reading of Fagan's critiques on the two judgments, I abstract the following three principles

Constitutionalism and the Common and Customary Law' (2010) 26 *South African Journal on Human Rights* 403, 413, 462–467.

²⁵ J van der Walt 'A Special Relationship with Women' (2002) 65 *Journal of South African Law* 148, 156.

²⁶ J van der Walt 'Horizontal Application of Fundamental Rights and the Threshold of the Law in View of the Carmichele Saga' (2003) 19 *South African Journal on Human Rights* 517, 524.

²⁷ See the introduction to J van der Walt *The Horizontal Effect Revolution and the Question of Sovereignty* (2014) 1–33.

²⁸ This also appears to be the view of the court in *RH v DE* [2014] ZASCA 133, 2014 (6) SA 436 (SCA). I critique this case in Zitzke 2015 (note 1 above). The stance of the SCA on common-law development expressed in *RH v DE* was overturned on appeal. See *DE v RH* [2015] ZACC 18 at paras 16–21.

²⁹ In A Fagan 'The Secondary Role of the Spirit, Purport and Objects of the Bill of Rights in the Common Law's Development' (2010) 127 *South African Law Journal* 611, 621–622, he clearly indicates his support for the fact that the common law could be developed on constitutional grounds: either because a right in the Bill of Rights requires development (per s 8 of the Constitution), because the interests of justice so require (per s 173 of the Constitution) or because the common law requires the development (per s 39(2) of the Constitution). Furthermore, Fagan's view is that if the common law is to be developed on one of the above three grounds, s 39(2) of the Constitution should kick in and the spirit, purport and objects of the Bill of Rights must be promoted.

³⁰ *Carmichele* (note 17 above) is critiqued in A Fagan 'Reconsidering Carmichele' (2008) 125 *South African Law Journal* 659.

³¹ [2005] ZACC 8, 2005 (6) SA 419 (CC), 2005 (9) BCLR 835 (CC), [2005] 8 BLLR 749 (CC) ('*K v Minister of Safety and Security*') is critiqued in A Fagan 'The Confusions of K' (2009) 126 *South African Law Journal* 159.

summarising his assessment as it is relevant for this discussion.³² Firstly, not all rules are developed whenever they are applied.³³ Secondly, the Constitution should only play a role in the developmental process and does not feature in the pure (non-developmental) application of the common law.³⁴ Thirdly, the Constitution does not impose duties on state functionaries – it only imposes duties on the state represented by the relevant Ministers – but even if the Constitution does impose duties on state functionaries, it would be unfair to hold state employees bound to constitutional obligations while non-state employees are not.³⁵ Abridging these summative points, Fagan is saying: the Constitution will be (and perhaps should be) an unnecessary consideration in most delictual matters. Being constitutionally heedless will, following Fagan, be the normal approach to dealing with delictual issues.

Johan Scott, even though celebratory of the Constitution's effect on cases relating to state negligence,³⁶ has recently critiqued (what he calls) the equitisation of the common law's development in cases where only non-state actors are involved.³⁷ His argument is that the Constitution has a devastating effect on private law as common-law development has the potential to disrupt a predictable set of rules that are necessary for purposes of legal certainty which in turn leads to effective commercial planning and strategising. Scott's claim is ultimately that the Constitution could be useful in delictual cases against the state but the invocation of the supreme law of South Africa could be problematic in all other cases. In cases involving non-state parties *inter se*, constitutional heedlessness would not be a bad thing in Scott's eyes.

As stated above, Neethling and Potgieter appear to form part of a more constitutionally optimistic paradigm. The duo indicates in their delict textbook that it should be accepted that the common law is in line with the Constitution unless the opposite is clearly apparent. They reason that there is a presumption in favour of constitutional compliance of the common law because the rights recognised in the Constitution are supported by the rights recognised at common law.³⁸ Even though this stance is significant, it is clear that the professors do not intend to complicate common-law reasoning with an approach that places constitutional scrutiny at the heart of every delictual dispute. Their approach is

³² This summary is inescapably incomplete as every line in Fagan's work contains a point of substance. However, for purposes of this piece the short condensation will have to do.

³³ Fagan (note 31 above) at 187, 190.

³⁴ See the subtext in Fagan (note 29 above) at 621ff and Fagan (note 31 above) at 178ff.

³⁵ Fagan (note 30 above) at 664–671 and Fagan (note 31 above) at 192. See also the similar argument made by S Wagner 'K v Minister of Safety and Security and the Increasingly Blurred Line Between Personal and Vicarious Liability' (2008) 125 *South African Law Journal* 673.

³⁶ See, eg, TJ Scott 'Vicarious Liability for Intentional Delicts – The Constitutional Factor Clinches Liability' (2013) 2 *Journal of South African Law* 348, 361; and TJ Scott 'Staatsaanspreeklikheid vir Opsetsdelikte van die Polisie – Die Hoogste Hof van Appél Kry Nogmaals Bloedneus' (2012) 3 *Journal of South African Law* 541, 557.

³⁷ TJ Scott 'Delictual Liability for Adultery: A Healthy Remedy's Road to Perdition' in JM Potgieter, J Knobel & RM Jansen (eds) *Essays in Honour of / Huldigingsbundel vir Johann Neethling* (2015) 421, 434. See also his concerns with regard to the equitisation of private property law in TJ Scott 'Effect of the Destruction of a Dwelling on the Personal Servitude of Habitatio' (2011) 74 *Journal of Contemporary Roman-Dutch Law* 155 especially at 168ff.

³⁸ Neethling & Potgieter (note 23 above).

to be welcomed insofar as they illuminate the theoretical compatibility of the common law and the Constitution. However, to the extent that they desire a business-as-usual approach, I distance myself from their line of enquiry because such an approach would amount to a failure to heed to the Constitution in most delictual disputes.³⁹

Also writing from a position of constitutional enthusiasm, Max Loubser and Rob Midgley dedicate a record 11 pages of their delict textbook to the interaction between delict and the Constitution.⁴⁰ What is interesting to note is that despite the fact that they would like to take the Constitution seriously, they do not provide much guidance as to when exactly the Constitution should 'actively' be considered in delictual disputes.⁴¹ It would appear that they favour a necessity test. When it would be necessary, is a question that is left to mystic, judicial intuition.⁴² Furthermore, the Constitution plays no apparent role in their 'systematic approach to delictual problem solving'.⁴³ Again, the approach of these authors is not as constitutionally heedful as it perhaps could be.

Despite the intricacies of each scholar's argument detailed above, the rudimentary common thread in their work is that the Constitution should not and/or will not have a substantive role to play in most delictual disputes, because, it seems, natural law in its human rights form is not all that important for the transformation of the common law, or the transformation of the common law is itself unnecessary. The implied support for constitutional heedlessness in all of these scholars' work leaves one wondering to what extent the larger-than-life monument of President Kruger, with its 'forceful presence' and 'air of steadfast resolution' that 'embodies the authority of ... political dominance' is then still being visited with admiration today.⁴⁴

To summarise my contention thus far: the general trend of aggression towards natural law in South Africa stems from the early 20th century. It was originally directed against the application of natural law to legislation, and is today directed against its application to the common law. It should be clear that I regard the natural law tradition as being foundational to both the Roman-Dutch part of our common law and human rights. I further regard natural law in its human rights form as an important development that should transform common law to keep it alive – 'keep alive' not only in the sense of having legal validity, but social validity too.⁴⁵ The merger of common law and human rights is foundational to a transformative theory and methodology for the South African common law.

³⁹ Similar concerns are raised against Neethling and Potgieter's stance by D Davis 'Legal Transformation and Legal Education: Congruence of Conflict?' (2015) *Acta Juridica* 172, 186ff.

⁴⁰ M Loubser & JR Midgley et al *The Law of Delict in South Africa* (2nd Edition, 2012) Chapter 2.

⁴¹ Ibid at 35.

⁴² I take note that these authors (at 34–35) rely on the case of *S v Thebus and Another* [2003] ZACC 12, 2003 (6) SA 505 (CC) at para 28 (Held that the criminal-law doctrine of 'common purpose' is constitutional) in forming their argument. See however the critique of judges being given the scope to consider the Constitution in whichever cases they like by Davis & Klare (note 24 above) at 464.

⁴³ Loubser & Midgley et al (note 40 above) at 23–26.

⁴⁴ The quoted phrases are derived from the description of the statue by P Labuschagne 'Memorial Complexity and Political Change: Paul Kruger's Statue's Political Travels Through Space and Time' (2011) 26 *South African Journal of Art History* 142, 145.

⁴⁵ C Douzinas & A Geary *Critical Jurisprudence* (2004) 18.

I turn to consider reasons that justify the application of a transformative theory and method for the development of the common law in general, and delictual disputes in particular. There are three such reasons, which also indicate why constitutional heedlessness is undesirable.

B Reasons Justifying a Rejection of Constitutional Heedlessness

1 Africanist Legitimacy

This first reason has two legs. First, the infiltration of an Africanist conception of human rights into the common law is important to ensure the legitimacy of the common law. Second, if the Africanist conception of rights is to be taken seriously, an extensive horizontal application of human rights must be fundamental to that enterprise. As to the first leg, the common law, fundamentally ‘white customary law’, was imposed on the South African legal system by conquest and has become the universal (ie ‘automatically applicable’) law in South Africa.⁴⁶ On the other hand, for any other type of customary law to be applied by a court, a whole host of requirements for its application need to be proven by litigants.⁴⁷ In a country where the majority of the population is not white, it is strange to imagine voluntary complicity in this state of affairs. I would speculate that issues of legal certainty and the closely related issues of national and transnational commercial stability probably played a key role in the decision taken during the negotiations for South Africa’s transition in the early 1990s to retain common law as a source of universal law insofar as it is consistent with the Constitution.⁴⁸ The inference that I draw from this negotiated position (which is a settlement somewhere in between a complete endorsement and rejection of the common law) is that the common law can remain legitimate in South Africa only if it is subject to a continuous constitutional audit so that a ‘new’ and ever-evolving South African common law can be established incrementally. Only this can justify the common law’s universal application.⁴⁹ If one accepts that the common law of South Africa fits quite comfortably in the classical liberal segment of natural legal thought, one might be tempted to argue that the reconciliation of constitutional rights and the common law is a superfluous endeavour because of the shared philosophical foundation between the two. However, the South African notion of

⁴⁶ Van Niekerk (note 15 above) at 21.

⁴⁷ Section 1 of the Law of Evidence Amendment Act 45 of 1988 allows courts to take judicial notice of customary law as long as it is readily ascertainable, sufficiently certain and not in conflict with the principles of public policy or natural justice. Evidence may be lead to prove the content of the customary law rule in question. The same caveats do not necessarily apply to the common law, which is assumed to be ascertainable and certain (even though a great deal of uncertainty still exists about the precise definitional components of the common law, see Van Niekers *ibid* at 21) and already imbued with the principles of public policy and natural justice as I have demonstrated earlier in this piece.

⁴⁸ Section 39(3) provides that ‘[t]he Bill of Rights does not deny the existence of any other rights or freedoms that are recognised or conferred by common law ... to the extent that they are consistent with the Bill’. Item 2(1) of Schedule 6 to the Constitution provides that ‘[a]ll law that was in force when the new Constitution took effect continues in force subject to (a) amendment or repeal; and (b) consistency with the new Constitution’.

⁴⁹ Davis & Klare (note 24 above) at 426 stress that the mission of the development clauses in the Constitution is to carry out an ‘audit and re-invention of the common law’.

constitutional rights differs in some respects from the classical liberal paradigm and for that reason has been referred to as being 'post-liberal'.⁵⁰ One of the most important differences is that the Africanist notion of human rights envisages a communitarian definition of human dignity.⁵¹ It is communitarian in the sense that the African Charter of Human and Peoples' Rights (Banjul Charter) 1981 is the only regional human rights instrument that explicitly and actively imposes duties on individuals to respect and protect rights of other individuals,⁵² which manifests in the Constitution as ss 8 and 39(2). This horizontal application of the Constitution is a key feature of post-liberal constitutionalism. The imposition of duties on non-state actors is significant because it demonstrates a concern for the values of 'cooperation, interdependence and collective responsibility'⁵³ as opposed to the individualistic ring to dominant Western notions of human dignity.⁵⁴ It may be that the concern with humane duties and mutual respect is a necessary check on the common law to ensure its legitimacy in 'post'-apartheid South Africa that was and is in such desperate need of reconciliation. One of the core aims of the democratic transition was to prevent South Africans from continuously turning a blind eye towards both 'privatised' and 'public' injustices.

The sense of duty promoted in the Africanist notion of human rights gives rise to the second leg of the reason under discussion. That is that horizontally applicable human rights need to be properly appreciated in order for the Africanist version of human rights to be given manifested validity by upholding a spirit of solidarity, generosity, unity and cohesion in South African common law.⁵⁵

2 *Deconstructive Substantive Equality*

Another aspect of the Africanist conception of human rights that is post-liberal is the acknowledgement of substantive equality as a legally genuine virtue.⁵⁶ The horizontal application of the Bill of Rights is important for purposes of recognising substantive equality in the South African context because it opens up the possibility for courts and other people who work with law to address the racist, patriarchal and economically oppressive effects of colonialism, apartheid and neo-colonialism.⁵⁷ In other words, horizontality opens up the possibility to

⁵⁰ K Klare 'Legal Culture and Transformative Constitutionalism' (1998) 14 *South African Journal on Human Rights* 146, 151, 153–156 (Argues that the South African Constitution breaks free from the classical idea of liberal constitutionalism and is 'post-liberal' because of its concern with social rights, substantive equality, positive state duties, horizontality, multi-culturalism, historical self-consciousness, and participatory governance).

⁵¹ J Cobbah 'African Values and the Human Rights Debate: An African Perspective' (1987) 9 *Human Rights Quarterly* 309, 324.

⁵² Chapter II (articles 27–29) of the African Charter. See also M Mutua 'The Banjul Charter and the African Cultural Fingerprint: An Evaluation of the Language of Duties' (1995) 35 *Virginia Journal of International Law* 339, 340, 359ff.

⁵³ Cobbah (note 51 above) at 320.

⁵⁴ Ibid at 324.

⁵⁵ D Chirwa 'In Search of Philosophical Justifications and Suitable Models for the Horizontal Application of Human Rights' (2008) 8 *African Human Rights Law Journal* 294, 303.

⁵⁶ Mutua (note 52 above) at 353.

⁵⁷ See J van der Walt 'Blixen's Difference: Horizontal Application of Fundamental Rights and the Resistance to Neo-Colonialism' (2003) 2 *Journal of South African Law* 311.

‘deconstruct’ (or to ‘map and critique’) the law.⁵⁸ It allows for mapping the law in that the entire body of law can be carefully re-examined and re-imagined in accordance with a new vision of social justice.⁵⁹ Closely related to the issue of mapping, I have previously argued that the reluctance to engage with substantive constitutional provisions in the development of common law tends to create the false (evolutionary functionalist) impression that common law develops along an objective, politically neutral path. Instead, constitutional attentiveness in the development of common law could force judges to acknowledge the political and ideologically contestable nature of decisions whether and how to develop the law.⁶⁰ Horizontality also creates a ‘legal’ mouthpiece for critiquing the law because it minimises the public-private divide that Marxists, feminists, queer theorists and critical race theorists argue serves to maintain various power imbalances in society – power imbalances that the transformative Constitution aims to substantively equalise.⁶¹ Individuals need to respect each other in their ‘private’ dealings with one another and the law should accommodate that respect and sense of duty that stems from a transformed vision of legal morality.⁶² Deconstruction as mapping and critique in this context, at first glance, seems to be contrary to legal certainty in a way that causes discomfort to some scholars. However, Dennis Davis and Karl Klare have lamented that a transformative theory for common law is ‘attentive to the values of stability, predictability and administrability’ because there will be many cases where the common law is constitutionally fine as it stands for the particular facts of a particular case.⁶³ However, common-law solutions are not timeless. They should always be subject to ‘reconsideration and contestation as experience progresses, understanding deepens, and/or circumstances change’.⁶⁴ This is the crux of a transformative theory for common law.

3 *The Single System of Law*

This last reason is inspired by André van der Walt’s interpretation of the often quoted extract from *Pharmaceutical Manufacturers* to the effect that there is one system of law in the democratic South Africa:⁶⁵

There is only one system of law. It is shaped by the Constitution which is the supreme law, and all law, including the common law, derives its force from the Constitution and is subject to constitutional control.

⁵⁸ C Albertyn & D Davis ‘Legal Realism, Transformation and the Legacy of Dugard’ (2010) 26 *South African Journal on Human Rights* 118, 205.

⁵⁹ H Cheadle & D Davis ‘The Application of the 1996 Constitution in the Private Sphere’ (1997) 13 *South African Journal on Human Rights* 44.

⁶⁰ This is the argument I first made in Zitzke 2016 (note 1 above).

⁶¹ Cheadle & Davis (note 59 above) at 45 and Chirwa (note 55 above) at 300–302.

⁶² Davis & Klare (note 24 above) at 411.

⁶³ *Ibid* at 412.

⁶⁴ *Ibid*.

⁶⁵ Van der Walt has written extensively on this topic, but his theory on the single system of law features most prominently in his book *Property and Constitution* (2012) 19–112 where he quotes and analyses the implications of *Pharmaceutical Manufacturers Association of South Africa and Another: In re Ex Parte President of the Republic of South Africa and Others* [2000] ZACC 1, 2000 (2) SA 674, 2000 (3) BCLR 241 para 44 (Held that the President’s decision to bring an Act of parliament into operation must be objectively rational).

In other words, common law (just like any other source of law) is not to be divorced from the Constitution. I argue that if the supremacy of the Constitution is to be taken seriously and judges are expected to properly justify their decisions whether to accept or alter prevailing common-law rules, then the Constitution should substantively feature in all common law disputes, whether it be to justify the prevailing rule or to develop it.⁶⁶ Davis and Klare similarly contend that a transformative method to common-law problems would not necessarily involve a complete rewrite of the common law in each case. All that should be required is for a lawyer to seriously and earnestly contemplate, at the start of each case, what potential constitutional provisions could influence the common law at stake in the dispute. Complimentarily, Van der Walt understands the single-system-of-law principle to mean that the common law needs to promote the spirit, purport and objects of the Bill of Rights in a way that advances both vested rights (whether derived from common law, legislation or the Bill of Rights) and the transformative goals of the Constitution.⁶⁷ If all law, including the common law, sings the same song (albeit sometimes in harmony and not in a monotone manner), the result is a single system of law.⁶⁸ From the discussion thus far, it should be clear that the single-system-of-law principle and its concern with realising substantive constitutional rights in the context of private common law can only be brought to fruition if a new transformative method is employed whereby all common-law disputes are constitutionally framed.⁶⁹

The implication of the above three reasons is that constitutional heedlessness is an approach that stifles the transformative project of, (a) affording legitimacy to common law through the incorporation of Africanist human rights jurisprudence into it; (b) deconstructing common law through mapping and critique; and (c) promoting the single-system-of-law principle that has been developed by the Constitutional Court to advance the supremacy of the Constitution. However, constitutional heedlessness still appears to be prevalent in various academic writings as I have detailed above. I now turn to demonstrate how constitutional heedlessness also features in recent judicial pronouncements.

C The Problem of Constitutional Heedlessness Manifested in Case Law

As explained in the introduction above, constitutional heedlessness involves a circumvention of the potential impact of the Constitution on the common law in a specific matter, even though the Constitution should play a role in that case. However, constitutional heedlessness does not involve an express rejection of the Constitution's potential impact. In other words, the Constitution is side-stepped by following a traditional, business-as-usual approach to dealing with the common law in a specific matter. At the same time, the court deciding a case or the commentator on a specific issue does not go out of their way to fight off the

⁶⁶ Zitzke 2015 (note 1 above) at 480.

⁶⁷ Van der Walt (note 65 above) at 20–21.

⁶⁸ Ibid at 26.

⁶⁹ Davis & Klare (note 24 above) at 412. See also the ideas developed by D Bhana 'The Role of Judicial Method in the Relinquishing of Constitutional Rights Through Contract' (2008) 24 *South African Journal on Human Rights* 300, 303 in the context of the common law of contract.

Constitution or explicitly push it aside. If the court or commentator did that, they would be employing an anti-constitutional approach.⁷⁰ What we are dealing with in cases of constitutional heedlessness is therefore simply neglecting to take the Constitution seriously in common-law matters. After reflecting on the decisions of *Loureiro SCA*, *Foetal Assessment Centre HC* and *Country Cloud SCA*, it could be argued that these cases all demonstrate the approach of constitutional heedlessness. In the following discussion, I intend to show that the Constitutional Court has on appeal responded to these cases in a way that resists the constitutionally heedless approach of the courts below.

1 *Loureiro SCA* & *CC*

In *Loureiro SCA*, the court had to determine whether a security company could be held contractually and/or delictually liable for the conduct of its security guard.⁷¹ The security guard had opened the Loureiro household's gate for a person who pretended to be a police officer while in reality the person was a robber. The robber then let his accomplices onto the property causing a great deal of financial and emotional harm to the Loureiro family and their employees.⁷² Writing for the majority, Mhlantla JA addressed the issue of the guard's negligence as well as the wrongfulness of his conduct.

On negligence, the court repeated the classical test articulated in *Kruger v Coetzee* that requires a court to determine whether 'a reasonable person in the position of the defendant would foresee the reasonable possibility of his conduct injuring another in his person or property and causing him patrimonial loss; and would take reasonable steps to guard against such occurrence; and the defendant failed to take such steps'.⁷³ Drawing from a variety of earlier cases, the court emphasised that the reasonable person is a normal, balanced individual,⁷⁴ and that the enquiry into reasonable foreseeability is an abstract enquiry where at least the general manner of the occurrence of harm should be anticipatable.⁷⁵ Furthermore, the reasonable person is not a prophet and therefore the determination of negligence should not be conducted 'wise after the event' – one must have regard to the specific circumstances that the guard found himself in.⁷⁶ Applying this test to the facts, the court held that the guard 'could not be faulted' for his assumption that the robber was a policeman because the robber arrived in a car with a blue flashing light and was dressed like a genuine police officer. There was no reason for the guard to have suspected the disguised persons of being robbers.⁷⁷ In a nutshell, the reasonable person in the guard's position 'would not have foreseen

⁷⁰ See Zitzke 2015 (note 1 above) at 470–472.

⁷¹ *Loureiro SCA* (note 4 above) at para 7. For purposes of this discussion, I will direct my attention to the delictual enquiry only.

⁷² *Ibid* at paras 4–6.

⁷³ *Kruger v Coetzee* 1966 (2) SA 428 (A) at 430E–G quoted in *Loureiro SCA* (note 4 above) at para 24.

⁷⁴ *Loureiro SCA* (note 4 above) at para 25.

⁷⁵ *Ibid* at para 26.

⁷⁶ *Ibid* at para 28.

⁷⁷ *Ibid* at paras 28–29.

that he was opening the gate to robbers and that he would be overpowered' and consequently the guard was not legally 'blameworthy'.⁷⁸

On wrongfulness the court held that it had to determine whether the guard breached a legal duty owed towards the Loureiro family. It found that the considerations applicable to determining negligence are the same considerations that apply in determining wrongfulness.⁷⁹ The equating of wrongfulness with negligence is peculiar because the court then proceeded to indicate that even if conduct is negligent it does not mean that it is wrongful.⁸⁰ The court then indicated that even though the determination of wrongfulness may involve taking the foreseeability of harm into account, wrongfulness primarily involves an enquiry into whether the legal convictions of the community require that the plaintiff be compensated for her/his losses caused by the negligence of the defendant.⁸¹ In this case the guard in question was under an obligation not to resist a policeman's entry to the property and, because he acted in good faith at all times believing that the robber was in fact a policeman, it cannot be said that the guard acted wrongfully.⁸² As a result, the security company was not held delictually liable for the guard's conduct because he acted neither negligently nor wrongfully.⁸³

On appeal, the Constitutional Court overturned the decision of the SCA partly because of the SCA's misunderstanding of the common law, partly because of a different reading of the facts and partly because of the SCA's failure to properly engage with the Constitution, or, as I like to formulate it, because of the SCA's constitutional heedlessness. Van der Westhuizen J, writing for a unanimous Court in *Loureiro CC*, framed the issue differently to the SCA. The judgment starts with a restatement that human dignity, the advancement of fundamental rights and the rule of law are the foundational values of the Constitution. It also found that the rights to life, freedom and security of the person, privacy and property were relevant to this case and needed to be protected.⁸⁴ On the basis of this transformative framework, the Court found that wrongfulness and negligence are two separate elements in the law of delict that should not be conflated. The SCA appeared to conflate the two elements and therefore its understanding of

⁷⁸ Ibid at paras 29–30.

⁷⁹ Ibid at para 31.

⁸⁰ Ibid at para 32.

⁸¹ Ibid at paras 32–33.

⁸² Ibid at paras 33–34.

⁸³ Ibid at para 35. The minority judgment portrays the opposite conclusion. Cloete JA held that the guard was negligent as he should have taken further steps to ascertain whether the robber truly was a police officer because the guard was a trained security professional (at para 49). Furthermore, the minority favoured a clearer separation of the elements of negligence and wrongfulness and stressed that the wrongfulness enquiry involves an engagement with constitutional norms that inform the legal convictions of the community to determine whether liability should be imposed on the defendant. The state of mind of the defendant is not relevant to the determination of wrongfulness (at paras 52–53). Because the guard's conduct is a positive act that infringed on the rights of the plaintiffs, it is well-established in the law of delict that wrongfulness is presumed in cases such as these. Therefore, Cloete JA would have held the security company delictually liable (at para 53).

⁸⁴ *Loureiro CC* (note 7 above) at para 1. Sections 12, 14 and 25 protect those respective rights. See also S Woolman & M Bishop 'Freedom and Security of the Person', D McQuoid-Mason 'Privacy', & T Roux 'Property' in S Woolman & M Bishop (eds) *Constitutional Law of South Africa* (2nd Edition, OS, 2008).

the common law was incorrect.⁸⁵ Wrongfulness zooms in on the conduct of the defendant and asks whether the legal convictions of the community, which are necessarily informed and shaped by the Constitution, regard the conduct as acceptable. Drawing from earlier judgments by the Constitutional Court and the SCA, Van der Westhuizen J noted that an analysis of wrongfulness should always involve constitutional contemplation. Failure to give due regard to the Constitution in determining wrongfulness could lead to a successful appeal on constitutional grounds.⁸⁶ It is clear that the SCA did not give proper consideration to constitutional imperatives in its decision on the wrongfulness of the guard's conduct. Furthermore, wrongfulness is based on three pillars: the duty to respect another's rights, the duty not to cause harm, and the reasonableness of imposing liability.⁸⁷ In the wrongfulness enquiry, the defendant's state of mind is not the focal point. The subjective state of mind of the defendant is the concern of the negligence enquiry that centres around the question whether the reasonable person *in the same situation* would have done the same.⁸⁸

Turning to the question of wrongfulness first, the Court held that the legal convictions of the community in this case was that security guards should not give criminals access to the properties that they are supposed to protect.⁸⁹ The test for wrongfulness is 'objective'⁹⁰ and thus the Court reasoned that liability should be imposed here because the constitutional rights to 'personal safety' and 'protection from theft or damage to property' deserve protection from security companies that are contracted to prevent the type of harm seen here.⁹¹ It appears that the infringement of constitutional rights justified a finding of wrongfulness without further ado. Thus, due to the fact that the Constitutional Court paid attention to the Constitution and the SCA did not, the SCA finding on wrongfulness was overturned. I will evaluate the Court's constitutionally enthusiastic approach to determining wrongfulness more thoroughly in Part III.

Moving on to the question of negligence, the Court repeated the test laid down in *Kruger v Coetzee* that the SCA also relied on.⁹² However, a different conclusion was reached on appeal to the Constitutional Court. Even though it is indicated that the test for negligence is partly normative and partly factual, it seems that the Court finds that the guard was negligent here on a different reading of the facts in the sense that certain facts that were not stressed in the SCA are emphasised here. The facts that indicate negligence are that the robbers arrived in an unmarked car, that the robber posing as a policeman wore a blazer (that

⁸⁵ *Loureiro CC* (note 7 above) at para 53.

⁸⁶ *Ibid* at para 34.

⁸⁷ *Ibid* at para 53.

⁸⁸ *Ibid* at para 53 (my emphasis). The negligence enquiry is partly subjective and partly objective: the test is objective insofar as we ask what the reasonable person would have done, but the test is subjectivised to suit the particular circumstances in which the defendant finds herself/himself.

⁸⁹ *Ibid* at para 55.

⁹⁰ Cf A Fagan 'Rethinking Wrongfulness in the Law of Delict' (2005) *South African Law Journal* 90, 92ff (Shows that the orthodox approach to wrongfulness, which defines it as an *ex post facto* (objective) enquiry, is not absolute because wrongfulness is sometimes determined with reference to *ex ante* considerations).

⁹¹ *Loureiro CC* (note 7 above) at para 56.

⁹² *Ibid* at para 58.

South African police officers do not wear while on duty), and that the policeman never announced the purpose for his visit and that he flashed his 'identity card' so quickly that the guard could make no proper evaluation of it.⁹³ The guard ought to have foreseen the possibility that robbers would want to gain access to the property by posing to be someone that they are not.⁹⁴ Answering the question of whether the reasonable person would have taken reasonable steps to prevent harm, the Court indicated that the extent of risk and consequences of the conduct was huge and that it would not have been disproportionately burdensome to have expected the guard to have taken reasonable steps to confirm the identity of the alleged policeman, to check that the policeman had lawful grounds to enter the property and, to attempt to make contact with his employer to obtain permission to allow the person onto the property.⁹⁵ The Court also underscored the fact that the guard in question was an A-grade security official. In cases where a person professes to have a certain level of skill, the 'greater the general level of expected care and skill will be'.⁹⁶ Evidently, the difference in the way that the facts are described by the Court is the main reason why the finding of negligence was made in the affirmative on appeal. Even though the Court is not clear on this, it would further appear that the SCA's failure to have due regard to the common-law rule of *imperitia culpa adnumeratur* ('lack of skill equals negligence') also contributed to its incorrect finding.⁹⁷ The failure to consider the *imperitia* rule has normative implications in that this rule promotes the notion that persons who are supposed to have better skills than others should be treated differently. This notion is consistent with the South African constitutional jurisprudence on the achievement of substantive equality that accentuates the need to treat different people with different characteristics differently to ensure that the playing field, even between private individuals *inter se*, is equalised.⁹⁸ In order for a court to take the constitutional audit of the common law seriously I argue that it should not just be open to constitutional redefinition of common-law rules, but it should also be open to justifying why extant common-law rules that do not need development are constitutionally compliant as they stand. If we do not openly justify why a common-rule is acceptable in its current form, it could lead to a type of constitutional heedlessness. In one sense, the SCA failed to consider a relevant common-law rule in its determination of negligence. In another sense, we could argue that the SCA's reasoning was also constitutionally heedless because of its failure to reflect on the implications of substantive equality for the negligence enquiry.

For these factual, common-law technical and constitutional differences, the decision of the SCA was overturned and the security company was held delictually liable on these facts. Importantly, for this discussion, at least one of the reasons

⁹³ Ibid at para 59–60.

⁹⁴ Ibid at para 61.

⁹⁵ Ibid at para 63.

⁹⁶ Ibid at para 64.

⁹⁷ See the observation in this regard made in TJ Scott's case note '*Loureiro and Others v iMvula Quality Protection (Pty) Ltd* 2014 3 SA (SCA)' (2014) 47 *De Jure* 374, 390.

⁹⁸ Section 9 of the Constitution. See also C Albertyn & B Goldblatt 'Equality' in S Woolman & M Bishop (eds) *Constitutional Law of South Africa* (2nd Edition, OS, 2008) Chapter 35.

why the SCA decision was wrong, was because of its constitutional heedlessness. Also important is the fact that in the Constitutional Court's judgment one can observe subtextual support for the single-system-of-law principle detailed earlier in this piece.

2 *Foetal Assessment Centre HC & CC*

In *Foetal Assessment Centre HC*, a mother represented her child who was born with Down syndrome, claiming damages suffered by the child due to the negligence of the Foetal Assessment Centre. The negligence was alleged to have been the failure of the Centre to have identified a high risk of abnormality in the foetus and to inform the mother of that risk which would have resulted in her terminating her pregnancy rather than letting the child be born.⁹⁹ The Centre took exception to the claim of the mother, alleging, among other reasons, that the claim is bad in law or contrary to public policy.¹⁰⁰ The question that Baartman J was confronted with was whether South African law could recognise a claim based on 'wrongful life'.¹⁰¹

A 'wrongful life' claim is brought by a child against a medical practitioner for the negligent misinformation communicated to the parents of the child about the risks of the pregnancy, resulting in the child being born (instead of being aborted) and suffering as a result of life with a disability. A claim of this nature should be distinguished from claims for 'wrongful pregnancy', that are brought by the parents of an unwanted but healthy child who would not have been born but for the medical practitioner's negligence (for example where a botched sterilisation is executed or where contraceptives are inadequately prescribed to parents who consult the medical practitioner with the aim of preventing pregnancy), as well as claims for 'wrongful birth' that are brought by the parents of a child born with certain congenital defects who would not have been born if the parents were properly informed of the risks involved with the pregnancy as they would have aborted the foetus. Claims for wrongful pregnancy and wrongful birth are recognised in South African law.¹⁰²

However, in the present matter, the court relied on the decisions in *Friedman v Glicksman*,¹⁰³ and *Stewart & Another v Botha & Another*,¹⁰⁴ to conclude that claims for wrongful life are not and should not be recognised in South African law. The cardinal reason for this decision and its predecessors is based on the concern that children with disabilities should not be told that their lives are 'wrongful'. This main concern can be expanded into four other closely related reasons. First, it would be contrary to the legal convictions of the community for a court to hold that children with disabilities would have been better off if they had

⁹⁹ *Foetal Assessment Centre HC* (note 5 above) at para 1.

¹⁰⁰ *Ibid* at para 4.

¹⁰¹ *Ibid* at para 5.

¹⁰² *Ibid* para 7.

¹⁰³ 1996 (1) SA 1134 (W) ('*Friedman*') 1142-1143 referred to in *Foetal Assessment Centre HC* (note 5 above) at para 9.

¹⁰⁴ [2008] ZASCA 84, 2008 (6) SA 310 (SCA), 2009 (4) All SA 487 (SCA) ('*Stewart*') at para 28 referred to in *Foetal Assessment Centre HC* (note 5 above) at para 20.

not been alive than to have ‘the unquantified blessing of life’.¹⁰⁵ Second, there can be no quantification of damage in comparing a position of existence and non-existence.¹⁰⁶ Third, a number of foreign jurisdictions have done away with claims for wrongful life and South Africa should follow this trend.¹⁰⁷ Fourth, the determination of wrongfulness in a case such as this questions whether it would have been better for the child not to have been born at all and that ‘goes so deeply to the heart of what it is to be human that it should not even be asked of the law’.¹⁰⁸ Even though counsel for the child in *Foetal Assessment Centre HC* contended that the constitutional rights of the child had not been considered in *Friedman* and *Stewart*, Baartman J concluded that there had not been a change in the legal convictions of the community since those decisions. This conclusion is finally backed up with the observation that many people with disabilities display great resilience and often overcome the odds of their condition, meaning that their lives cannot be ‘wrongful’ and therefore the exception was upheld.¹⁰⁹ I contend that the circumvention of the potential impact of the Constitution in the determination of wrongfulness in this case is reflective of constitutional heedlessness that I have been describing throughout this piece. This is so because even though it was argued that the Constitution was not taken seriously in earlier decisions on the topic of wrongful life, Baartman J was committed to approaching the common law as if it was business as usual, circumventing the issue of constitutional application and reimagination. Simply assuming that the common law on a specific topic and the Constitution are harmonious without deeper engagement is symptomatic of constitutional heedlessness.

On appeal to the Constitutional Court, Froneman J, writing for a unanimous Court, held that there were two problematic parts to the High Court decision. First, the exception was readily granted. There could potentially be a claim for the child based on the facts and the High Court was perhaps too quick to uphold the exception irrespective of the prevailing common law rules possibly needing development. Second, the High Court failed to properly contemplate whether its decision pertaining to the wrongfulness of the Centre was truly reconcilable with constitutional rights and values, including the best interests of the child standard that is guaranteed in s 28 of the Constitution.¹¹⁰ Despite these problematic aspects in the High Court decision, the Constitutional Court only provided a new framework within which the High Court would have to reconsider the matter.¹¹¹

With regard to the first challenge, Froneman J indicated that in order for an exception to succeed, there should be no possible reading of the facts that could give rise to a cause of action.¹¹² If the possibility for the development of common law arises, it may be best to refuse the granting of exceptions or orders for absolution from the instance. This is especially true where there are complex

¹⁰⁵ *Foetal Assessment Centre HC* (note 5 above) at para 9.

¹⁰⁶ *Ibid.*

¹⁰⁷ *Ibid* at para 19.

¹⁰⁸ *Ibid* at para 20.

¹⁰⁹ *Ibid* at para 29.

¹¹⁰ *Fetal Assessment Centre CC* (note 7 above) at para 81.

¹¹¹ *Ibid* at para 81.

¹¹² *Ibid* at para 10.

factual matrices with uncertain legal positions accompanying those facts, even though this is not a hard and fast rule.¹¹³ In a case such as the present one where a common-law rule could be changed altogether, it would usually be wise to refuse the exception so that all of the evidence and arguments could be heard for an informed decision to be made about whether or not the relevant common-law rule should be developed.¹¹⁴ As indicated above, the reason why the Constitution should play a central role in all common-law matters is to promote the single-system-of-law principle. This important principle was missed in both problematic aspects of the High Court judgment. The *Fetal Assessment Centre CC* judgment highlights the single-system-of-law principle with great enthusiasm. Froneman J reiterated that the development clauses in the Constitution have the aim of ensuring that constitutional values permeate the common law.¹¹⁵ Thus, both in its failure to consider the possibility of development and constitutional compliance in general, the High Court fell short of its transformative mandate.

With regard to the second challenge, the Constitutional Court showed that the term ‘wrongful life’ is an incorrect reflection of what a claim of that nature really involves. A claim for ‘wrongful life’ does not truly involve labelling the life of the child as being wrongful. The claim involves determining whether ‘the law should allow a child to claim compensation for a life with a disability’.¹¹⁶ By framing the issue in that way, the enquiry focuses on the fact that the law cannot ignore the difficulties that a child born with a disability is faced with. The dictum that has historically been repeated by our courts, that the law should not determine an essential question that seeks to define what it means to be human, is not acceptable in a single system of law where the Constitution is supreme. By side-stepping this question, judges attempt to exempt themselves from making a difficult value choice. They only ‘attempt to’ circumvent the value choice, because deciding not to answer the question has practical implications that in themselves display a particular value choice that is disguised in a fictitious cloak of neutrality.¹¹⁷ Thus, the decision that the child in this case should have no claim has a practical, value-laden consequence: children with disabilities deserve no special treatment or legal protection, despite the difficulties that they may face. Moreover, there cannot be areas of life and law where the Constitution can simply be ignored. In other words, the question about whether a claim for so-called wrongful life should be recognised by our law must be answered in light of the Constitution. It is not an extra-legal issue.¹¹⁸ The question then arises: what should the influence of the Constitution be on this part of the law?

Foreign law may be useful in coming to an answer.¹¹⁹ Even though there are jurisdictions that do not recognise claims of this nature, there are jurisdictions that do. Different jurisdictions often have different answers to the same legal

¹¹³ Ibid at paras 11–12.

¹¹⁴ Ibid at para 24.

¹¹⁵ Ibid at para 14 referring to *K v Minister of Safety and Security* (note 31 above) at paras 16–17.

¹¹⁶ *Fetal Assessment Centre CC* (note 7 above) at para 19.

¹¹⁷ Ibid at para 22.

¹¹⁸ Ibid at para 23.

¹¹⁹ Ibid at para 28. Section 39(1)(c) of the Constitution provides that a court may consider foreign law when interpreting the Bill of Rights.

question because of differing socio-political circumstances and contexts. The task that a court faces therefore is to decide which jurisdictions have similar normative frameworks and material contexts to our own. Phrased differently, the exercise of employing a comparativist method involves questioning whether our objective, constitutional, normative framework could draw substance from the foreign jurisdiction in question.¹²⁰ In this matter, foreign jurisdictions that emphasise the best interests of children and the autonomy of parents would probably be compatible jurisdictions. This is a transformative approach to legal comparativism that compliments the single-system-of-law principle. The High Court's cursory reliance on foreign law is therefore an undesirable treatment of that source of law.

The Court identified the rights to equality, dignity and the best interests of children as relevant to the issue at hand.¹²¹ Even though common-law rules can often be easily interpreted to be harmonious with the Constitution, there are cases such as the present one where the rules do not, as they stand, optimally promote all of the relevant constitutional provisions. The current common-law model does not give due regard to the need to assist persons with disabilities to realise their right to be substantively equal to other people – especially not for children who have the right to have their best interests considered paramount in every case relating to them.¹²² This would especially be true in cases where parents do not pursue a claim for wrongful birth and the child is then left without a remedy.¹²³ Ultimately, the Court tacitly endorses the notion that the Constitution has an important, deconstructive role to play in ensuring substantive equality between non-state actors. Furthermore, the child's dignity is not optimised by denying their claim in the circumstances of this case. Even though the common-law position may appear to create the impression that life with a disability is equally worth living than life without a disability, awarding the child the right to claim in these circumstances would be more sensitive to the child's condition that may require extra resources to live comfortably.¹²⁴ In conclusion, the Constitutional Court held that the High Court erred insofar as it upheld the exception without appropriately considering whether the relevant common-law rule needed to be developed and the 'factual, legal and policy issues' that should have been established to play a decisive role in the court's decision.¹²⁵ Even though Froneman J did not make the final decision on whether the common law had to be developed in this case, it is clear that the constitutional heedlessness of the High Court was seriously questioned and is not to be repeated in similar matters in future.

¹²⁰ *Fetal Assessment Centre CC* (note 7 above) at paras 32 and 42.

¹²¹ *Ibid* at para 49. Sections 9, 10 and 28 of the Constitution guarantee those respective rights. On the right to equality, see Albertyn & Goldblatt (note 98 above). See also S Woolman 'Dignity' & A Friedman, A Pantazis & A Skelton 'Children's Rights' in S Woolman & M Bishop (eds) *Constitutional Law of South Africa* (2nd Edition, OS, 2008).

¹²² *Fetal Assessment Centre CC* (note 7 above) at para 59.

¹²³ *Ibid* at paras 61–62.

¹²⁴ *Ibid* at para 72.

¹²⁵ *Ibid* at para 78.

3 *Country Cloud SCA & CC*

In both *Loureiro CC* and *Fetal Assessment Centre CC* the constitutional heedlessness employed by the courts below was held to be unfitting and inappropriate on appeal. In those appeals to the Constitutional Court, a transformative method was employed where the Constitution played a central role in the understanding of the common law that resulted in the appeals being upheld. *Loureiro CC* and *Fetal Assessment Centre CC* both implicitly give support to the ideas that the common law derives its legitimacy from strong horizontal constitutional application, that the common law often needs deconstruction, and that the single-system-of-law principle is important. Due to the lack of constitutional lustre in the SCA and High Court judgments, those decisions were found to be substantively incorrect. On the other hand, in the matter of *Country Cloud* the constitutional heedlessness of the SCA did not substantively have a practical effect on the outcome of the case, as the SCA's decision was confirmed on appeal to the Constitutional Court. However, a more complete and analytically rigorous approach by the SCA to the issue at hand would have involved constitutional considerations, as was done by the Constitutional Court on appeal.

In *Country Cloud SCA*, the Department of Infrastructure Development in the Gauteng province contracted with a construction company called Ilima Projects for the erection of a clinic in Soweto. The Department undertook to pay R480 million to Ilima for the completion of the work. Assisted by the Department, Ilima entered into a loan agreement with Country Cloud Trading for R12 million in order to embark on the project.¹²⁶ After the loan was made available and paid to Ilima, the Department cancelled the building contract leading to Country Cloud suffering damages on account of Ilima being liquidated and the principal debt (plus interest) consequently not being repaid.¹²⁷ The SCA held that a valid contract had been entered into and that the cancellation of that contract had not been unlawful.¹²⁸ The question that had to be answered was whether the Department wrongfully caused Country Cloud's pure economic loss on these facts.

After surveying the history of the common-law position on the causing of pure economic loss in the law of delict, Brand JA, writing for a unanimous SCA, explained that the element of wrongfulness in delict acts as a 'safety valve' to prevent limitless liability.¹²⁹ Wrongfulness is determined with reference to the legal convictions of the community and questions the reasonableness of imposing liability on the defendant in accordance with public policy.¹³⁰ There had been no case with similar facts that a court has had to decide in the past and thus the court had to resolve whether or not the common law had to be developed to allow Country Cloud's claim here.¹³¹ Brand JA held that, even though a blameworthy state of mind and foreseeability of harm are relevant policy considerations to the

¹²⁶ *Country Cloud SCA* (note 6 above) at paras 1 and 5.

¹²⁷ *Ibid* at para 2.

¹²⁸ *Ibid* at paras 15–16.

¹²⁹ *Ibid* at paras 17–18.

¹³⁰ *Ibid* at paras 19–20.

¹³¹ *Ibid* at para 26.

determination of wrongfulness in cases of pure economic loss, if the question here turned on whether the Department foresaw the possibility of harm and whether the Department intentionally proceeded in its harm-causing conduct, then the result would be indeterminate liability. This is so because a long list of third parties (including Ilima's employees and other creditors) would then be able to claim from the pure economic loss caused by the Department's cancellation of the contract.¹³² The court reasoned that that would be an undesirable state of affairs and that in determining who should bear the loss in cases such as these, the doctrine of 'vulnerability to risk' should be employed. That doctrine dictates that if a defendant could reasonably have protected themselves against the risk that materialised, then the defendant should bear the risk.¹³³ Applied to the facts of this case, Country Cloud could have either claimed repayment of the money that it lent to Ilima or it could have taken cession of Ilima's claim against the Department. Because no substantial reasons could be provided as to why Country Cloud did not take these steps to protect itself against the risk that materialised, the Department could not be said to have acted wrongfully towards Country Cloud and so there was no delictual liability in that case.¹³⁴ Even though there is ample authority that shows that the determination of wrongfulness should involve constitutional considerations, the SCA opted to circumvent constitutional considerations here. Constitutional heedlessness won again.

On appeal to the Constitutional Court, the only issue that had to be addressed was whether the Department acted wrongfully towards Country Cloud in cancelling the contract. Khampepe J, for a unanimous Court, reiterated the thorough overview on the law of wrongfulness that the SCA had provided with some variations and one key added ingredient: relying on its earlier decision in *Loureiro CC*,¹³⁵ the Court emphasised the fact that the legal convictions of the community, that shape the element of wrongfulness, had to be constitutionally understood.¹³⁶ Considerations relating to the blameworthy state of mind of the alleged wrongdoer, the prevention of indeterminate liability and the vulnerability to risk doctrine are indeed relevant policy considerations to determining wrongfulness.¹³⁷ In addition to these considerations the constitutional value of state accountability should, at least, be considered.¹³⁸

Section 1(d) of the Constitution affirms accountability as a founding value of the democratic state.¹³⁹ The value of state accountability could, but will not always, be translated into a private-law duty that finds delictual liability.¹⁴⁰ In previous

¹³² Ibid at para 28.

¹³³ Ibid at para 30.

¹³⁴ Ibid at paras 31–33.

¹³⁵ *Loureiro CC* (note 7 above) at para 53.

¹³⁶ *Country Cloud CC* (note 8 above) at para 21.

¹³⁷ Ibid at paras 39–43 and 51–61.

¹³⁸ I take cognisance of the fact that Country Cloud raised the argument based on state accountability in the Constitutional Court and that the Court did not raise this issue *mero motu*. For a sound overview of the different ways by which the value of state accountability can be realised see A Price 'State Liability and Accountability' (2015) *Acta Juridica* 313.

¹³⁹ See further T Roux 'Democracy' in S Woolman & M Bishop (eds) *Constitutional Law of South Africa* (2nd Edition, OS, 2008).

¹⁴⁰ *Country Cloud CC* (note 8 above) at para 45.

cases, state accountability translated into a private-law duty only where the state functionaries concerned acted maliciously for personal gain either through corrupt, fraudulent or otherwise criminal conduct.¹⁴¹ The state functionaries in this case did not act illicitly towards Country Cloud. The only wrong that the state committed here was against Ilima who could hold the state accountable by instituting a claim based on their contract.¹⁴² Upholding a claim in favour of Ilima based on contract and a claim in favour of Country Cloud based on delict would undermine the functioning of the Department concerned by depleting its resources. That is a relevant consideration in promoting state accountability because the state cannot be accountable to the public if courts undermine its functions.¹⁴³ For this and other reasons (that were slight variations on the same themes present in the SCA judgment) the appeal was dismissed.¹⁴⁴

III CONSTITUTIONAL OVER-EXCITEMENT

A The Problem of Constitutional Over-Excitement in *Loureiro CC*

In the previous part, I sought to make a case for the rejection of constitutional heedlessness as an approach to common-law issues. I demonstrated that there is no insurmountable conceptual or jurisprudential barrier that insulates common law from the influence of human rights. I argued that it is desirable for the common law to be infused with constitutional norms for the purposes of ensuring the common law's legitimacy in light of Africanist notions of human rights, that the much needed transformation of private law could be guided by the Constitution's development clauses that aim to map and critique the common law and, that the single-system-of-law principle developed by the Constitutional Court requires that the Constitution be taken seriously even in seemingly uncontroversial issues. The approach that I promote can be broadly referred to as a transformative theory for common law, even as applied between non-state actors. The vision for a transformed common law is presently supported neither in the delict scholarship nor the recent High Court and SCA judgments surveyed above. However, recent Constitutional Court jurisprudence rejects the approach of constitutional heedlessness as a legitimate private law method that certain academics and courts appear to support.

The question left to answer is to what extent the Constitution should take possession of the common law. On the one hand, one could argue that the Constitution should completely dispose of well-established common-law rules and that private law should be completely rewritten in every case. This approach

¹⁴¹ Ibid at paras 46–47.

¹⁴² Ibid at para 50.

¹⁴³ Ibid.

¹⁴⁴ Ibid at para 69.

I call ‘constitutional over-excitement’.¹⁴⁵ On the other hand, one could follow a moderate yet transformative approach that tries to solve tensions between various sources that may potentially apply to any given case. In this part, I show why it may be said that the wrongfulness enquiry in *Loureiro CC* might have bordered on constitutional over-excitement and why that is undesirable.

To recap, in *Loureiro CC*, the Court held that conduct is wrongful if the legal convictions of the community, constitutionally understood, regard it as unacceptable. ‘It is based on a duty not to cause harm – indeed to respect rights – and questions the reasonableness of imposing liability.’¹⁴⁶ Here the Court determined wrongfulness with special attention given to the duty to respect rights. But not any rights: the constitutional rights to personal safety and property that should be protected by security companies and their guards who are contracted for that purpose.¹⁴⁷ Other than the condensed definition of wrongfulness, the fact that it should not be conflated with negligence and that wrongfulness should be determined before negligence, very little doctrinal discussion about wrongfulness is endeavoured and previous similar cases were not visited to fit this case into an interesting and complex pattern of the wrongfulness theme. Indeed, it seems that wrongfulness is simply (and outwardly uncontroversially) established whenever constitutional rights are infringed. Thus, it appears that wrongfulness can be established by investigating constitutional law alone, with no need to consider long-standing precedent. This is a point that Alistair Price takes issue with in his note on *Loureiro CC*.¹⁴⁸ Even though Price agrees with the outcome of the case, he takes issue with the reasoning of the Court.

First, Price argues that the Court emphasised the open-ended policy considerations and underemphasised the importance of ‘principled analogy from past or hypothetical cases where legal duties in delict have been or would be imposed or denied’.¹⁴⁹ This type of reasoning, Price contends, facilitates the ‘orderly and incremental development of the common law’ that ensures a greater degree of ‘coherence and predictability’.¹⁵⁰ Price is clear on the fact that by this he does not mean that the previous decisions and their principles provide unconditional demands, as analogical reasoning is also complemented by policy considerations.¹⁵¹ What he finds problematic, in a quasi-Dworkinian fashion, is that principles should not be replaced by constitutional policy considerations. It would appear that Price is in favour of finding a balance between and integration

¹⁴⁵ The fact that the Constitutional Court sometimes gets the common law of delict wrong and relies too strongly on the Constitution is not in itself a new idea. The phrase that I employ here aims to provide a generic term for a common problem. See, eg, Fagan on the *Carmichele* judgment (note 30 above), on *K v Minister of Safety and Security* (note 31 above) and most recently also in ‘Causation in the Constitutional Court: *Lee v Minister of Correctional Services*’ (2014) 5 *Constitutional Court Review* 104. Fagan’s enthusiasm about the Constitution’s potential for the law of delict has been curbed by these judgments.

¹⁴⁶ *Loureiro CC* (note 7 above) at para 53.

¹⁴⁷ Sections 12 and 25 of the Constitution guarantee these rights.

¹⁴⁸ A Price ‘The Contract/Delict Interface in the Constitutional Court’ (2014) 25 *Stellenbosch Law Review* 501.

¹⁴⁹ *Ibid* at 503–504.

¹⁵⁰ *Ibid* at 504.

¹⁵¹ *Ibid*.

of common-law principles and constitutional policy.¹⁵² Ultimately, for Price, the reasoning of the Court would have been more complete, more analytically rigorous and therefore more defensible if it referred more extensively to previous similar decisions.

Price draws our attention to the facts of *Viv's Tippers (Edms) Bpk v Pha Phama Staff Services (Edms) Bpk*¹⁵³ that are analogous to those in *Loureiro CC*, but the conclusions of the respective cases seem to be at odds.¹⁵⁴ In *Viv's Tippers* a security company was sued because its employee failed to prevent thieves, masquerading as mechanics, from stealing a vehicle. The legal snag was that the owner of the vehicle left it on the property of the party with whom the security company had a contract, but the owner did not have any agreement with the security company. In *Loureiro CC*, the plaintiff's family members who were harmed also did not have a contract with the security company. The SCA in *Viv's Tippers* held that the security company owed no duty grounded in delict towards the owner of the vehicles. In *Viv's Tippers* the SCA overturned *Compass Motors Industries (Pty) Ltd v Callguard (Pty) Ltd*,¹⁵⁵ and doubted the correctness of *Longueira v Securitas of South Africa (Pty) Ltd*.¹⁵⁶ Price is concerned that these conflicting cases received no mention in *Loureiro CC* despite the facts in all these matters being consonant. Uncertainty therefore now exists with regard to whether *Viv's Tippers* is wrong and *Compass Motors* and *Longueira* are correct, or whether all of these cases are distinguishable from one another.¹⁵⁷ Thus, the Court failed to properly grapple with the common law and instead opted to decide the matter simply on the basis of a breach of constitutional rights. Thankfully Price provides us with useful insight into why *Viv's Tippers* and *Loureiro CC* are reconcilable. For Price, the two cases are distinguishable primarily because *Viv's Tippers* related to the situation where a security guard omitted to prevent theft while *Loureiro CC* related to a security guard who positively acted by opening a gate for robbers. The difference in the description of the conduct in each case is important, because wrongfulness is presumed where negligent positive conduct causes physical harm, while negligent omissions are *prima facie* lawful.¹⁵⁸ These nuances in the law of wrongfulness were not properly addressed in *Loureiro CC* and consequently it seems as though the Constitution has come to replace the technicalities of the common law. If the Court wrestled with analogous precedent in *Loureiro CC*, it would have contributed to establishing a more complete picture on the different scenarios that could play out where a security company commits (or does not commit) a delict.¹⁵⁹ Below I elaborate on why we might agree with Price's scepticism of the constitutional over-excitement in *Loureiro CC*.

¹⁵² Ibid at 505.

¹⁵³ [2010] ZASCA 26, 2010 (4) 455 (SCA) ('*Viv's Tippers*').

¹⁵⁴ Price (note 148 above) at 505.

¹⁵⁵ 1990 (2) SA 520 (W) ('*Compass Motors*').

¹⁵⁶ 1998 (4) SA 258 (W) ('*Longueira*').

¹⁵⁷ Price (note 148 above) at 506.

¹⁵⁸ Ibid at 507.

¹⁵⁹ Ibid at 510.

B Reasons Justifying a Rejection of Constitutional Over-Excitement

1 Transformation is not Revolution

The basic tenor of Price's view should be favoured. Principled reasoning in common-law matters is not necessarily untransformative.¹⁶⁰ The implication of *Carmichele* and similar matters is that courts do not only need a thorough knowledge of the Constitution in order for incremental developments to be effected to the common law – they also need a thorough knowledge of common law in order for its developmental exercise to be meaningful and well-reasoned.¹⁶¹ To state it differently, it is crucial to know what is inside of the box before rejecting the box off the cuff. Otherwise you cannot be sure that you are truly thinking outside of the box. In a similar vein, Klare is clear about the fact that transformative legal reasoning goes beyond visionless conservation and reform, but nevertheless falls short of a revolution.¹⁶² A complete displacement of common law rules, without properly knowing what they are, coupled with a sole reliance on the Constitution, sounds a lot more like a revolution than a transformation. Earlier I have said that the Constitution has an important role to play in securing the legitimacy of common law, the deconstruction of common law, and securing the realisation of the single-system-of-law principle. It is important to note that it is *common law* that needs to be legitimised and deconstructed – we actually have to work with the common law and take it seriously to do these things. If we throw the common law out completely, there is nothing to deconstruct or legitimise. Furthermore, the single-system-of-law principle can be secured by permeating common law with constitutional spirit. We do not have to throw the entire common law out and replace it with the Constitution to ensure a single legal system. Going to the extreme of ignoring the common law is, perhaps surprisingly to many people, just as untransformative as neglecting the Constitution in private-law disputes. It would certainly be revolutionary and possibly even decolonial to do away with the common law as a whole without further ado. But that is simply not what the Constitution requires.

2 The Constitution also has its Limits

Price's critique of the over-emphasis on constitutional principles is also valid because it guards against what Lourens du Plessis would refer to as a purely monumental reading of the Constitution whereby the Constitution is celebrated and regarded as the pinnacle of an already transformed society,¹⁶³ without the critical recognition that the Constitution as a source of law is inherently limited

¹⁶⁰ See, eg, T Roux 'Transformative Constitutionalism and the Best Interpretation of the South African Constitution: Distinction without a Difference' (2009) 20 *Stellenbosch Law Review* 258.

¹⁶¹ *Carmichele* (note 17 above) at para 40 implies that a thorough knowledge of the common law is necessary to determine whether the common law as it stands is consonant with the Constitution.

¹⁶² Klare (note 50 above) at 150.

¹⁶³ L du Plessis 'The South African Constitution as Monumental and Memorial and the Commemoration of the Dead' in B Pieroth & R Christensen (eds) *Rechtstheorie in Rechtspraktischer Absicht* (2008) 189, 192–193.

in its transformative capacity.¹⁶⁴ The Constitution should not be deified as a perfect source of law because it might not hold the answers and solutions to all of South Africa's questions and problems. For example, Sanele Sibanda contends that the Constitution has not solved the poverty and spatial justice problems in South Africa.¹⁶⁵ True social and economic transformation can only be realised if complimented by something in addition to the law and clever judgments, such as strong social movements and/or activist politics.¹⁶⁶ Pius Langa also says that one of the biggest barriers to social transformation relates to the failure on the part of the beneficiaries of colonialism and apartheid to create a climate suitable for reconciliation, which cannot truly be rectified by the creation of any law, including the Constitution. This does not necessarily mean that the beneficiaries must be punished severely for their privilege but that the beneficiaries must at least play an active role in the process of reconciliation by making contributions towards building a South Africa united in our diversity.¹⁶⁷ In short, the Constitution has its limits and constitutional over-excitement fails to take cognisance of those limits; constitutional over-excitement perhaps expects too much from the Constitution. To avoid the monumentalisation of the Constitution, Du Plessis argues for a simultaneous monumental and memorial reading of the Constitution that does not over-celebrate nor under-appreciate the significance of the supreme law. To put it in my terms, the approach called for should not be constitutionally heedless, nor over-excited. The midway between the two extremes is best achieved, with specific reference to dealing with the common law, through the doctrine of what Du Plessis calls 'adjudicative subsidiarity' that guards against constitutional absolutism while simultaneously having due respect for the Constitution's supremacy.¹⁶⁸

IV ADJUDICATIVE SUBSIDIARITY AS MIDWAY BETWEEN THE EXTREMES

Adjudicative subsidiarity refers to the 'reading strategy'¹⁶⁹ that the Constitutional Court has employed in the past to ensure that the Constitution would not be 'overused',¹⁷⁰ subject to the caveat that the supremacy of the Constitution must

¹⁶⁴ For a good overview on the South African literature regarding the inherent limitations of the Constitution's transformative aspirations, see D Brand *Courts, Socio-Economic Rights and Transformative Politics* (2009) (unpublished LL.D Thesis, Stellenbosch University) 3ff.

¹⁶⁵ S Sibanda 'Not Purpose-Made! Transformative Constitutionalism, Post-Independence Constitutionalism and the Struggle to Eradicate Poverty' (2011) 22 *Stellenbosch Law Review* 482, 488ff. Cf F Michelman 'Liberal Constitutionalism, Property Rights, and the Assault on Poverty' (2011) 22 *Stellenbosch Law Review* 706 (Disagrees with Sibanda's contention that the Constitution is 'too liberal' to deal with poverty). See also Sibanda's reply to Michelman in S Sibanda 'Not Quite a Rejoinder: Some Thoughts and Reflections on Michelman's "Liberal Constitutionalism, Property Rights and the Assault on Poverty"' (2013) 24 *Stellenbosch Law Review* 329.

¹⁶⁶ P Langa 'Transformative Constitutionalism' (2006) 17 *Stellenbosch Law Review* 351, 358.

¹⁶⁷ Ibid at 359.

¹⁶⁸ Du Plessis (note 163 above) at 197. See also his other works on adjudicative subsidiarity: L du Plessis 'Subsidiarity: What's in the Name for Constitutional Interpretation and Adjudication?' (2006) 17 *Stellenbosch Law Review* 207 ('Du Plessis on Subsidiarity') that formed the basis of L du Plessis 'Interpretation' in S Woolman & M Bishop (eds) *Constitutional Law of South Africa* (2nd Edition, OS, 2008) Chapter 32, 150–158.

¹⁶⁹ Du Plessis on Subsidiarity (note 168 above) at 209.

¹⁷⁰ Ibid at 215.

be taken seriously.¹⁷¹ Van der Walt conceptualises subsidiarity as a reconciliation of the dictum in *S v Mblungu and Others*, that it is ‘a general principle that where it is possible to decide any case, civil or criminal, without reaching a constitutional issue, that is the course which should be followed’¹⁷² and the single-system-of-law principle enunciated in *Pharmaceutical Manufacturers*.¹⁷³ The fact that subsidiarity is a midway between two extreme approaches to constitutional application must be emphasised. On the one hand, subsidiarity is not a greenlight for constitutional heedlessness or a mechanism for purifying the common law from politics. (Due to the fact that subsidiarity is attentive to the Constitution’s call for a single system of law, it is ‘a politics confirming and -enhancing device that ensures interplay between constitutional principles and democratic laws, reformist initiatives and vested rights, change and stability’.)¹⁷⁴ On the other hand, subsidiarity fights against constitutional over-excitement in that it requires lawyers to take legislation, common law and customary law seriously, in light of and subject to the Constitution, in a way that allows a multiplicity of legal sources to peacefully coexist without complete methodological chaos. In the context of property law, Van der Walt says that the main purpose of subsidiarity is to structure the ‘choice of the source of law’.¹⁷⁵ In other words, subsidiarity provides ‘guidelines that identify the source of law that primarily governs litigation’ related to rights infringements.¹⁷⁶ Practically, subsidiarity can be of use to the law of delict as well.

As a point of departure, it is helpful in all cases to start by identifying a constitutional right that has potentially been infringed by an alleged wrongdoer. From this point forward, Van der Walt provides us with two subsidiarity principles.

The first principle is derived from *South African National Defence Union v Minister of Defence and Others*.¹⁷⁷ In that case the Constitutional Court held that if a constitutional right is alleged to have been infringed, the dispute must be resolved in accordance with legislation that has specifically been promulgated to protect the right concerned. Thus, existing legislation cannot be thoughtlessly circumvented in favour of sole reliance on a constitutional right.¹⁷⁸ The rationale for this first principle seems to be that legislation of this kind gives content to a constitutional right and so there is no need to reinvent the wheel by placing sole reliance on the constitutional text. Legislation relevant to disputes relating to constitutional-right infringements that take on delictual form include, for example, the Road Accident Fund Act¹⁷⁹ and the Compensation for Occupational Injuries Act.¹⁸⁰ Both of these enactments protect the constitutional rights to dignity and bodily integrity.¹⁸¹ However, the fact that these enactments must be applied and taken

¹⁷¹ AJ van der Walt ‘Normative Pluralism and Anarchy: Reflections on the 2007 ‘Term’ (2008) 1 *Constitutional Court Review* 77, 90ff.

¹⁷² [1995] ZACC 4, 1995 (3) SA 867 (CC), 1995 (7) BCLR 793 (CC) at para 59.

¹⁷³ Van der Walt (note 171 above) at 100.

¹⁷⁴ Ibid.

¹⁷⁵ Van der Walt (note 65 above) at 35.

¹⁷⁶ Ibid.

¹⁷⁷ [2007] ZACC 10, 2007 (5) SA 400 (CC), 2007 (8) BCLR 863 (CC) at paras 51-52.

¹⁷⁸ Van der Walt (note 65 above) at 36.

¹⁷⁹ Act 56 of 1996.

¹⁸⁰ Act 30 of 1993.

¹⁸¹ See, eg, HB Klopper *The Law of Third-Party Compensation* (3rd Edition, 2012) 10ff.

seriously in the disputes that they regulate does not mean that the Constitution becomes completely irrelevant. In the process of interpreting legislation, s 39(2) of the Constitution kicks in and the spirit, purport and objects of the Bill of Rights must be promoted. Alternatively, one could attack the legislation for constitutional invalidity following s 172 of the Constitution. Practically, we end up with an amalgamation of legislation and constitutional spirit instead of a complete circumvention of the Constitution (constitutional heedlessness) or a complete circumvention of legislation (constitutional over-excitement).

The second principle is derived from *Bato Star Fishing (Pty) Ltd v Minister of Environmental Affairs and Others*.¹⁸² In that case the Constitutional Court held that if a constitutional right is alleged to have been infringed, the dispute must be resolved in accordance with legislation that has specifically been promulgated to protect the right concerned and existing legislation cannot be thoughtlessly circumvented in favour of sole reliance on common law.¹⁸³ Of course, this principle is subject to the proviso that if legislation does not cover the dispute in question, common and customary law act as legal safety nets to provide rules and principles to regulate the matter.¹⁸⁴ Davis and Klare are of the view that common and customary law are always being incrementally developed whenever those sources are used.¹⁸⁵ There certainly is critical merit to their argument. The consequence of Davis and Klare's view is that s 39(2) – which requires a court to promote the spirit, purport and objects of the Bill of Rights when the common or customary law is developed – is always relevant whenever the common or customary law is engaged. However, if we take a more sceptical view of Davis and Klare, like Fagan indirectly does,¹⁸⁶ common and customary law are not always being developed whenever those sources are adjudicated on. Sometimes, common or customary law is quite simply applied. In my view, that does not necessarily mean that the Constitution becomes irrelevant to cases where the common law is applied. Section 173 of the Constitution bestows on courts the power to develop the common law considering the interests of justice. That power could imply a choice between applying the law or developing it. The exercise of judicial power by making a choice between application and development should be properly justified in a transformative democracy to give effect to the rationality principle that is fundamental to the founding value of the rule of law.¹⁸⁷ The justification process must necessarily involve serious constitutional engagement because the common and customary law can only survive in our constitutional democracy if it is consonant with the Constitution, following ss 2, 39(3) and Item 2(1) of Schedule 6 to the Constitution. To be clear, I argue that even if a rule is to be applied with no substantive change, a sharp statement should be made as to

¹⁸² [2004] ZACC 15, 2004 (4) SA 490 (CC), 2004 (7) BCLR 687 (CC) at para 25.

¹⁸³ Van der Walt (note 65 above) at 36.

¹⁸⁴ Ibid.

¹⁸⁵ Davis & Klare (note 24 above) at 423–424.

¹⁸⁶ Fagan (note 31 above) at 187, 190.

¹⁸⁷ Section 1(c) of the Constitution stipulates the rule of law as one of the founding values of the South African state. For more on the 'culture of justification' in 'post'-apartheid South Africa, see E Mureinik 'A Bridge to Where? Introducing the Interim Bill of Rights' (1994) 10 *South African Journal on Human Rights* 31, 32.

why the rule is constitutionally compliant as it stands so that a court observes its justificatory mandate. Thus, the Constitution is always speaking in common and customary law matters, even if we accept that a difference exists between the application and development of those sources. Practically, we end up with an amalgamation of common or customary law and the Constitution, instead of a complete circumvention of the Constitution (constitutional heedlessness) or a complete circumvention of common law and legislation (constitutional over-excitement).

This scheme of working with various legal sources largely coincides with the methodology laid out in s 8(3) of the Constitution.¹⁸⁸

When applying a provision of the Bill of Rights to a natural or juristic person in terms of subsection (2), a court –

- (a) in order to give effect to a right in the Bill, must apply, or if necessary develop, the common law to the extent that legislation does not give effect to that right; and
- (b) may develop rules of the common law to limit the right, provided that the limitation is in accordance with section 36(1).

The method of s 8 also starts with an identification of a potentially relevant constitutional right, followed by a consideration of legislation that gives effect to that right. If no such legislation exists, the common law (and, following s 211(3), I would add ‘or customary law, if it is applicable’) is applied or developed to give effect to the relevant constitutional right. Yet, s 8 is only useful when non-state actors are engaged in a dispute. I argue that the two broad principles of subsidiarity advocated by Van der Walt, discussed above, provide us with a fallback for what to do when the state is an alleged wrongdoer and s 8 does not apply.

A problem with the method of s 8 and the broader principles of subsidiarity is that it does not provide a final shield against constitutional right infringements where no legislation, common or customary law are applicable. I venture to say that this is where s 38 of the Constitution provides a potential solution. That section grants the power to courts adjudicating Bill of Rights issues to provide ‘appropriate relief’ where rights are ‘infringed or threatened’.¹⁸⁹ On the front

¹⁸⁸ Woolman has for a long time been campaigning for a stronger reliance on s 8 or so-called ‘direct constitutional application’ that would lead to a more coherent and thorough constitutional rights jurisprudence. See, eg, S Woolman ‘Application’ in S Woolman & M Bishop (eds) *Constitutional Law of South Africa* (2nd Edition, OS, 2008) Ch 31; and S Woolman ‘The Amazing, Vanishing Bill of Rights’ (2007) 124 *South African Law Journal* 762. Many academics seem to have heeded Woolman’s call: See, eg, D Bhana ‘The Horizontal Application of the Bill of Rights: A Reconciliation of Sections 8 and 39 of the Constitution’ (2013) 29 *South African Journal on Human Rights* 351; D Davis ‘Where is the Map to Guide Common Law Development?’ (2014) 25 *Stellenbosch Law Review* 3; N Friedman ‘The South African Common Law and the Constitution: Revisiting Horizontality’ (2014) 30 *South African Journal on Human Rights* 63. Bhana, Davis and Friedman have in recent times also noted that s 8 of the Constitution should be the starting point for navigating constitutional application. In these works it seems that the distinction between direct and indirect constitutional application should be collapsed. The result is ultimately an endorsement of subsidiarity as I have described it here.

¹⁸⁹ See generally *Ramakatsa and Others v Magashule and Others* [2012] ZACC 31, 2013 (2) BCLR 202 (CC) (Held that s 19 of the Constitution can bind non-state actors without recourse being had to legislation or common law); M Dafele ‘The Directly Enforceable Constitution: Political Parties and the Horizontal Application of the Bill of Rights’ (2015) 31 *South African Journal on Human Rights* 56; and P de Vos ‘It’s My Party (And I’ll Do What I Want To)?: Internal Party Democracy and Section 19 of the South African Constitution’ (2015) 31 *South African Journal on Human Rights* 30.

of the law of damages, the possibility exists for an aggrieved party to claim constitutional damages if a constitutional right has been infringed, but legislation, common and customary law leaves them remediless. In *Fose v Minister of Safety and Security* the Constitutional Court held that a litigant cannot claim constitutional damages in addition to common-law damages grounded in delict where they have been assaulted by employees of the state.¹⁹⁰ Hidden in *Fose* is a vote of support for the principle of subsidiarity. If a constitutional right has been infringed and there is no legislation, the common law is relied on. Only if the common law then fails to provide adequate relief will a litigant be able to claim constitutional damages. Only once a litigant has reached the end of the 'sources rope' can constitutional damages be claimed.¹⁹¹

The methodological approach to sources that subsidiarity provides can assist in finding a midway between constitutional heedlessness and over-excitement. However, it is not a foolproof method. Subsidiarity can very easily be abused if too much focus is placed on avoiding the Constitution in favour of other sources. In order for subsidiarity to have true transformative flair, the appliers of subsidiarity must always be conscious of the dual-purpose philosophy underpinning it. That is, we need a single system of law while simultaneously being cautious of placing too large a burden on the Constitution at the expense of an integrated reading of various applicable sources. In critical spirit, I must further highlight that subsidiarity might not be the only approach to moderate constitutional heedlessness and over-excitement. In fact, subsidiarity itself might have to be approached with circumspection so that it does not become crystallised, closed or venerated in itself. In order to be a truly critical approach to the issue of constitutional application it must be self-reflective, subject to change and, if necessary, be open to deconstruction and reconstruction. This is so because, as the prolific mystic philosopher Rumi teaches us, once we believe that we have mastered something, we should run from that false state of finality and accomplishment.¹⁹²

V CONCLUSION

Approaches such as those championed by Price, Du Plessis and Van der Walt aim to strike a balance between the two extremes of constitutional avoidance (in this discussion, constitutional heedlessness) and constitutional over-excitement deserve support. Constitutional heedlessness is undesirable as it stifles the

¹⁹⁰ [1997] ZACC 6, 1997 (3) SA 786 (CC), 1997 (7) BCLR 851 (CC) ('*Fose*') paras 19–54.

¹⁹¹ This principle indirectly also featured in *President of the Republic of South Africa and Another v Modderklip Boerdery (Pty) Ltd* [2005] ZACC 5, 2005 (5) SA 3 (CC), 2005 (8) BCLR 786 (CC) (Constitutional damages payable by the state were granted to a farm owner whose constitutional property right had been infringed by unlawful occupiers who could not be evicted without their constitutional right to housing being infringed) and directly featured in *Minister of Police v Mboweni and Another* [2014] ZASCA 107, 2014 (6) SA 256 (SCA), [2014] 4 All SA 452 (SCA) (Held that constitutional damages cannot be claimed in loss-of-breadwinner cases where the law of delict provides sufficient protection to the rights of children whose parents pass away).

¹⁹² J Rumi *Masnavi i Ma'navi – Book Five* (trans K Helminski & C Helminski, 2000) 564–565: 'Know the true definition of yourself. That is essential. Then, when you know your own definition, flee from it.'

development of common law that is needed in order for it to maintain its legitimacy in the transformative South African legal era. Constitutional over-excitement also stifles the potential of forming a critical framework for evaluating the common law as it fails to realise that the Constitution should be approached unpretentiously regarding its limited transformative possibilities. Just as much as the Constitution is important, it is not a perfect tool to effect real and tangible change in the South African society. Such a humble recognition is crucial to monumentalise neither common law nor the Constitution.

Analogous to Paul Kruger's problematic metaphorical monument of legal reasoning, the emblematic monument of the Constitution might have a similar haunting effect on private common law reasoning. While these two monuments can be impressive and inspirational at first glance (and for a while after that), the modest recognition needs to be made that the required interplay between the common law and the Constitution was bargained and determined in a process of ideological negotiation and struggle where various parties to the discussion had to sacrifice certain beliefs regarding certain sources of law. Those sacrifices serve as a memorial to both the common law and the Constitution. Thus, a transformative method, inspired by the formation of the negotiated South African constitutional democracy, should be a sign of memorialising both sources of law. And perhaps, just perhaps, adjudicative subsidiarity may help us craft a unified memorial concurrently dedicated to the common law and the Constitution. With that said, the rise of decolonial theory might soon hit private law. When that happens, our memorial and everything that it represents could see itself covered in paint or it could even be on the brink of being completely dismantled from its pedestal.

Against the Interests of Justice: Ignoring Distributive Justice When Certifying Class Actions

*Khomotso Moshikaro**

I INTRODUCTION

Following the advent of our constitutional democracy, our procedural law must be applied to serve the interests of justice.¹ This is made clear by s 173 of the Constitution, which gives courts ‘the inherent power to protect and regulate their own process, and to develop the common law, taking into account the interests of justice’. One would expect, therefore, that since ‘justice’ is the compass guiding our procedural law’s application, there would be some attempt by our courts, most of all our Constitutional Court, to give clarity as to what the term ‘interests of justice’ means in a principled and coherent manner. This is not the case. Instead, court decisions about the interests of justice comprise mainly of vague ‘factors’ that apply on a ‘case by case’ basis depending on the ‘facts and circumstances’ of each case. At first glance, this approach seems imminently sensible since we are told that it ensures ‘flexibility’ in court process.

I shall argue that at second glance, this approach is fallacious. The Constitutional Court would do better, and is in fact required to adopt an approach that strives for clarity in any principle’s content and that sensibly defines the classes of cases that call for the application of certain procedural rules. I shall argue that this approach is demanded not only by the rule of law as some have argued,² but also by the form of ‘justice’ contemplated by the constitutional text – distributive justice. The need for clarity allows courts to make decisions on how to distribute scarce court resources according to moral judgments that subsidise good litigation planning by litigants. Before tackling the argument proper, a short note on methodology.

I shall address the fallacies underpinning this ‘flexible’ approach against the backdrop of the Constitutional Court’s judgment in *Mukaddam*,³ which dealt with the place of class actions in our law. *Mukaddam* will serve as a central case that allows us to examine this flexible approach in its best form. It is not only

* Lecturer, Department of Private Law, University of Cape Town. LLB (UP) BCL (Oxon) MSc in Contemporary Chinese Studies (Oxon). My thanks go to Anton Fagan, Helen Scott and Leo Boonzaier for their invaluable critiques and suggestions. Further acknowledgement goes to Ofentse Motlasedi and Louis Botha for their long sufferance of my monologues over the subject of this article.

¹ Section 173 of the Constitution of the Republic of South Africa, 1996. Underscored by *Chief Lesapo v North West Agricultural Bank and Another* [1999] ZACC 16, 2000 (1) SA 409, 1999 (1) BCLR 1420 (CC).

² S Woolman ‘The Amazing, Vanishing Bill of Rights’ (2007) 124 *South African Law Journal* 762.

³ *Mukaddam v Pioneer Foods (Pty) Ltd and Others* [2013] ZACC 23, 2013 (5) SA 89 (CC), 2013 (10) BCLR 1135 (CC) (*‘Mukaddam’*).

a landmark judgment on class actions, but is in many ways a polemic against ‘rigidity’ in our law.⁴

The argument proper focuses on three main fallacies that underlie the Court’s flexible approach. The first fallacy is that the interests of justice is an entirely flexible standard. The Court conceptualises the interests of justice in so vague and unprincipled a manner that the flexible approach becomes arbitrary, violating the legality principle. Others have criticised the Court’s vagueness along similar lines. How this article differs is that it flatly rejects certain myths about what is required when courts apply a legal principle according to ‘the facts and circumstances’ of a case. Further, it provides an explanation of what it means for a court to make a moral assessment that decides a case ‘on the facts’. Second, the Court assumes general flexibility overrides certification requirements. It fails to see that the proper place of flexibility in procedural law is when courts decide to *extend* an already existing and clear rule, or cause of action, which currently applies to a particular class of people, to another class that previously did not have the benefit of that rule’s protection. It is, therefore, a problem of distributive justice that requires courts to justify the rule’s extension. This appreciation of the distributive problem would give guidance to lower courts when these courts exercise their discretion to regulate their own process, rather than the ‘maybe/maybe not’ guidance afforded by the flexible approach. Third, the Court fallaciously believes predictability equals rigidity. This springs from the continued invocation of what I shall call the spectre of the rigid and ‘formalist’ judge, which is invoked erroneously as a solution in search of a problem, or conveniently when arguments against opponents have been exhausted and one is in desperate need of a useful *ad hominem* or straw-man.

II BACKGROUND

The *Mukaddam* decision came at the tail end of a long price-fixing saga. Initially, South Africa’s largest bread producers colluded to fix the price of bread and formed a cartel. The Competition Commission eventually discovered this anti-competitive practice and fined the defendants. Subsequently, a coalition of different NGOs and trade unions sued the producers for damages in the form of a class action on behalf of bread consumers. Concurrently, a coalition of bread distributors brought a separate class action suing for damages. This is the *Mukaddam* decision. Both the distributors and consumers failed to have their class actions certified in the High Court and appealed.

The Supreme Court of Appeal (SCA) delivered a lucid and detailed judgment that explored the requirements for a class action to be certified.⁵ These requirements according to Wallis JA are (i) the class must be defined; (ii) there must be a common claim or issue that can be determined by way of a class action; (iii) there must be a valid cause of action; (iv) the legal representative must be suitable to represent the members of the class; and (v) a court must be satisfied

⁴ Ibid at para 34, where Jafta J states categorically: ‘rigidity has no place in our law.’

⁵ *Trustees for the time being of Children’s Resource Centre Trust and Others v Pioneer Food (Pty) Ltd and Others* [2012] ZASCA 182, 2013 (2) SA 213 (SCA), 2013 (3) BCLR 279 (SCA).

that a class action is the most appropriate procedure to adopt for the adjudication of the underlying claims.

The SCA found that the consumer's litigation had a *prima facie* basis in law and remitted the claim to the High Court to consider further evidence before deciding on the issue. However, the SCA had little patience for the distributor's claim and in a terse judgment, Nugent JA declined to certify the class because the merits of the suit were at face-value unconvincing.⁶ Further, the SCA put weight to the fact that there are two types of class actions. Opt-in class actions are voluntary and require members of the class to willingly join the class to vindicate their rights. Opt-out class actions are initially involuntary because individuals form part of the class automatically. Nugent JA ruled that considering opt-in class actions are voluntary and akin to joinders, exceptional circumstances must be present to justify why the claim could not have been brought as a joinder instead of a class action.⁷

The distributors then approached the Constitutional Court. In a surprising judgment, the majority found in favour of the distributors. Jafta J found for the majority that the SCA had not applied the certification criteria in the interests of justice. In particular, the SCA apparently treated this criteria rigidly by considering them to be requirements or conditions precedent. Instead they are 'factors' that help a court determine where the interests of justice lie. There are supposedly instances when some of these criteria do not have to be met and yet still a court should certify a class action.⁸ The Court went so far as to refuse to say that a claim for constitutional damages based on a Bill of Rights violation may not require certification at all.⁹ This is apparently because s 38(a) of the Constitution grants class action recourse 'as of right'.¹⁰ Having outlined the factual matrix of the article, I now move to a critique of the Court's approach in *Mukaddam* based on the fallacies identified earlier.

⁶ *Mukaddam and Others v Pioneer Food (Pty) Ltd and Others* [2012] ZASCA 183, 2013 (2) SA 254 (SCA) ('*Mukaddam SCA*'). Nugent JA found that competition legislation did not protect profits and neither does the Constitution. He further found that joinders sufficiently deal with the concerns opt-in class actions aim to deal with, and so there needs to be exceptional circumstances present to justify certifying such class actions.

⁷ *Ibid* at para 14.

⁸ Jafta J states in *Mukaddam* (note 3 above) at para 35 that –

'these requirements must serve as factors to be taken into account in determining where the interests of justice lie in a particular case. They must not be treated as conditions precedent or jurisdictional facts which must be present before an application for certification may succeed. The absence of one or another requirement must not oblige a court to refuse certification where the interests of justice demand otherwise.'

⁹ The exception is a partial dissent by Mhlanhla AJ (as she was then) who expressed some confusion as to why certification would be barred by the constitutional text or would operate against the interests of a plaintiff. I explain below why Justice Mhlanhla was correct in her assessment, but that she conceded far too much in agreeing with the majority on their 'factor'-based approach.

¹⁰ *Mukaddam* (note 3 above) at para 40.

III FALLACY I: THE INTERESTS OF JUSTICE IS AN ENTIRELY FLEXIBLE STANDARD

A The Section 173 Power and Legality

In an article published almost ten years ago, Stu Woolman critiqued the Constitutional Court's Bill of Rights jurisprudence as evincing a violation of the rule of law. The main thrust of the article was that the Court in its reasoning ought to craft and follow intelligible rules in their interpretation of the Bill of Rights.¹¹ To state the obvious, courts lay down such 'rules' because their decisions do not only bind the parties before them, but set precedent for other courts faced with similar facts.¹² This reflects what Neil MacCormick has accurately called the 'universal' quality of legal reasoning.¹³ This universal quality is not just simply because of precedent, but because of the justification underpinning a particular decision.¹⁴ It is simply 'logically impossible' to ignore this underlying justification for one decision when considering a similar category of cases that follows.¹⁵

This article takes on a different challenge. The aim of this piece is to explore what happens when the Constitutional Court is enjoined to apply purely moral criteria and exercise judgement 'in the interest of justice'. The challenge is to understand the role of the rule of law, if any, and how courts pursue justice in the exercise of their discretionary control over their own process. The charge against the Constitutional Court is that it has failed to appreciate that no notion of justice can extricate itself from this universal quality and to do so is a violation of the rule of law and the notion of treating like cases alike. The first hurdle in advancing the rule of law objection to the Court's 'flexible' approach, therefore, is to demonstrate that the rule of law is a viable norm at all. The primary objection to this is exemplified by this statement by Drucilla Cornell:

The goal of a modern legal system is rational synchronization and not rational coherence. Synchronization recognizes that there are competing rights situations and real conflicts between the individual and the community which may not yield a 'coherent' whole. The

¹¹ Woolman (note 2 above) at 791. To use his words:

'I work within a tradition of constitutional law — of which South Africa is most avowedly a part — that recognizes rules as a necessary feature of the legal landscape. (Their ontological status, within that constitutional order, may well be contested.) The problem with which this paper is concerned is the extent to which our Constitutional Court fails to generate cognizable legal rules and meaningful precedent.'

¹² Justice Kate O'Regan may have put this better stating: 'As a member of the bench, I am required to issue a judgment; and in that judgment, I am required to lay down a rule of law that binds both the parties before the court and South African society as a whole.' (see S Woolman *ibid* at 789.) See also J Raz *The Authority of Law: Essays on Law and Morality* (2002) 213.

¹³ N MacCormick 'Legal Reasoning and Legal Theory' (1978) 78.

¹⁴ *Ibid* at 81. See also R Dworkin 'The Supreme Court Phallanx' (2007) *New York Review of Books* (2007) 92, 99 where the author argues that a court follows precedent out of 'respect not for the narrow holding of earlier cases, one by one, but for the principles that justify those decisions'.

¹⁵ For instance in a case of manufacturers liability for a defective product, a party may win a case because she is a *consumer* who is owed a duty by a manufacturer. Only her being a consumer, who suffered harm caused by a manufacturer's product, normatively justifies this award of damages. See Lord Tomlin's famous speech in *Donaghue v Stevenson* (1932) AC 562 at 599, 1932 SC (HL) 31 at 57 to illustrate this need for a normative category.

conflicts may be mediated and synchronized but not eradicated... . In reality, a complex, differentiated community can never be reduced to a single voice.¹⁶

The argument seems to be that some norms are irreducibly contested and cannot be resolved through legal reasoning. This view may be especially tempting concerning procedural law because procedural rules apply to a myriad of substantive issues from contract, to delict, to constitutional and all other sorts of claims. Therefore, the only choice open to courts is surely to take this ad hoc approach and avoid speaking with ‘a single voice’.

Hardly. The rule of law does not require seamless coherence. A primary purpose of the rule of law principle is to ensure predictability. The value of this predictability is that it allows individuals to plan their lives by anticipating state action. It does not necessarily follow, therefore, that in order to be predictable, all conflicts of legal rules are to be eliminated. Compliance with the rule of law is a matter of degree.¹⁷ The key is that legal rights cannot be contradictory with respect to the same class of cases and therefore incapable of guiding the behaviour of citizens. If the law’s aim is to guide behaviour, it cannot give people equal reasons to engage in behaviour that is mutually incompatible.¹⁸ This is directly self-defeating.¹⁹ However, there is no need to argue that a competing right or norm is ‘eradicated’ — we speak of better or stronger reasons to comply. Few, if no, advocates of the rule of law have ever sought the ‘seamless’ coherence Cornell speaks of. The rule of law has only ever sought to ensure laws are coherent *enough* (it does not have to be a coherent whole) to allow law to guide individual behaviour.

The point is that predictability obtains if persons can understand and conform to rules without further direction or clarification by officials.²⁰ Therefore, the degree of generality, or explanations given as to how one right trumps another, need only be clear enough to allow citizens to reasonably predict how certain legal conflicts would reasonably be decided were they to be challenged in court. As to procedural law, this means enough predictability for a lawyer to advise a client about what procedural requirements ought to be met to frame the merits of a case.

B The Nature of the Section 173 Power

Section 173 of the Constitution states that courts have an inherent power to protect and regulate their own process and to develop the common law on matters of procedure, consistently with the interests of justice. I shall focus on the process regulation power. In order to judge whether the Constitutional Court in *Mukaddam* has exercised this power properly, we must understand the nature of

¹⁶ D Cornell ‘Institutionalization of meaning, recollective imagination and the potential for transformative legal interpretation’ (1988) 136 *University of Pennsylvania Law Review* 1135, 1211.

¹⁷ Raz (note 12 above) at 215.

¹⁸ *Makbanya v University of Zululand* [2009] ZASCA 69, 2010 (1) SA 62 (SCA), [2009] 8 BLLR 721 (SCA) at paras 9 and 82 on why Nugent JA found it ‘striking’ that two ratios in the same judgment could be ‘mutually destructive’.

¹⁹ D Parfit *Reasons and Persons* (1984) 4.

²⁰ HLA Hart *Concept of Law* (1961) 207.

the power. This power has been described by the Court as a discretion.²¹ Because this power is discretionary, it is apparently possible for two different courts to have come to different reasonable conclusions on the same matter.²² In fact, the Court has even suggested that there may not be a single characterisation for the powers in s 173.²³ All this is rather puzzling since one expects judges who abide by the rule of law to guarantee some measure of predictability. If this discretion is so wide, how can potential litigants or the public predict how courts are to decide future cases? This seems to be an arbitrary power that flouts the rule of law.

Not so. In this section I aim to show how it is possible for a power that does not require judges to create legal rules through precedent can still be exercised in a manner that secures the predictability demanded by the rule of law. This is important because later, I will demonstrate how even with this lowered standard for predictability, the Constitutional Court in *Mukaddam* still failed to meet basic rule of law standards. The first step will be to appreciate that this power requires judges to make a moral judgement demanded by the facts and circumstances of a case. The important implication of this is that when judges do this, they do not create binding legal rules. However, this does not mean courts do not create principles that apply generally beyond the particular facts and circumstances of a case. We must therefore dispel several myths about the ‘relevant’, ‘specific’ or ‘particular’ nature of the facts and circumstances of a case. We must also understand what it conceptually means to decide a case ‘on the facts’. From this, we can understand how and why judges still comply with the rule of law when exercising this s 173 discretion, by exploring how applying moral judgements create guidelines that give judges reasons to look backwards at previous court decisions, even though they are not obliged to do so in the same manner they would if they were bound by legal precedent.

C Facts and Circumstances

I wish to first address the broad interpretation of the flexibility doctrine. By this I mean reading the constant stress the Court places on the facts and circumstances of a case as meaning that this requires a court to adopt what I shall call ‘itemised’ decision-making. This is the notion that the *uniqueness* of the facts and circumstances of each case must determine how a procedural rule applies. Jafta J may be read to make a similar assertion in *Mukaddam* when discussing the discretion of lower courts:

It is the court before which the class action is brought which is best placed to determine whether a class action *in relevant circumstances* will be in the interests of justice.²⁴

If we are uncharitable and interpret this passage to mean that the Court believes the circumstances of every case call for unique consideration of what the interest of justice standard demands as to whether a class action ought to be granted, then

²¹ *South African Broadcasting Corporation Ltd v National Director of Public Prosecutions and Others* [2006] ZACC 15, 2007 (1) SA 523 (CC), 2007 (2) BCLR 167 (CC) at para 38–40.

²² *Ibid.*

²³ *Ibid.*

²⁴ *Mukaddam* (note 3 above) at para 46.

the rule of law objection to this position is obvious.²⁵ Litigants who wish to bring a class action simply cannot predict by any measure whatsoever whether a court will grant such an action before incurring the costs of such litigation. But this answer is also not sufficient. What then do courts mean when they use the phrase ‘facts and circumstances’?

MacCormick suggests an answer. There are two kinds of facts in a case. There are *operative* facts that bring into operation a certain legal rule.²⁶ And there are inoperative facts that are simply normatively and legally irrelevant. This means that when judges say a case must depend on the facts and circumstances, this is not saying that the facts of each case are so unique that the basis of the very legal rule itself must be considered, but that determining the operative facts of a case is not often an exact science and requires judgement of some kind in deciding what qualifies as operative facts. But more to the point, this argument on uniqueness is simply untenable as a matter of logic. It would suppose that it is possible to have a good reason to decide a single case which is not a good generic reason for deciding cases of the particular type under consideration.²⁷ However, this bogey of uniqueness is not so easily slain. Courts sometimes argue we ought to decide cases according to their ‘particular facts and circumstances’. In ordinary language ‘particular’ is often synonymous with ‘specific’ and not necessarily ‘relevant’ things. Could this reference to particular circumstances then be distinct from deciding what ‘relevant’ or operative facts are?

Let us investigate. Courts sometimes refer to ‘particular’ facts and circumstances when distinguishing one case from another. This was exemplified in *Chief Lesapo*,²⁸ a case which the Court cited in *Mukaddam*. That case considered the constitutionality of a provision that prevented individuals approaching courts when their property was seized to recover certain debts. When considering a previous case that had found a similar provision barring judicial review was acceptable, Mokgoro J said:

However, the decision in the present case must be understood in the context of its *particular* circumstances, which differ from those of the revenue cases. We are not called upon to decide the correctness or otherwise of the conclusion in *Hindry* and we refrain from doing so.²⁹

Two points stand out here. First, Mokgoro J claims the facts of the case before her are so different from prior authority that the case requires a different approach. This is not controversial. Second, she claims it is the *particular* circumstances that demand this different approach. It is the use of the word ‘particular’ that I believe may confuse some advocates of ‘flexibility’ in decision-making.

²⁵ Fortunately, because of the Court’s penultimate paragraph in *Mukaddam* (note 3 above) at para 47, which tries to pull back a lot of its flexibility rhetoric, I do not believe that this is what the Constitutional Court meant. I address the argument only to better explain the latter part of this section regarding matters of fact and the general nature of moral, not just legal, reasoning.

²⁶ MacCormick (note 13 above) at 97.

²⁷ Ibid.

²⁸ *Chief Lesapo* (note 1 above). See *Mukaddam* (note 3 above) at para 30 for Jafta J’s reliance on *Chief Lesapo*.

²⁹ *Chief Lesapo* (note 1 above) 28 (emphasis added).

Legal theorists and judges have recognised that a certain legal principle may have different levels of generality in its scope and application.³⁰ Consider the issue of same-sex equality rights. At the lowest level of generality, a principle of equal protection or personal liberty may mean that same-sex relations must be decriminalised,³¹ whilst at a more general level it may include equal marriage rights.³² If when judges speak of deciding cases according to ‘particular’ facts, they mean deciding a case on the narrowest terms or at the lowest level of generality, I shall not take issue with this approach in this paper.³³

My quarry is of a different kind. In this section, I aim to dispel the myth that an exercise of discretion according to s 173 that takes account of facts and circumstances means an ad hoc discretion which grants license for courts to decide similar cases in any direction. Let us not be fooled by a higher court not interfering with a lower court’s discretion. It does not necessarily follow that this is because an exercise of the s 173 discretion is irreducibly contestable.³⁴ However, there is a possible argument for the notion that discretion over process is irreducibly contestable. This may be that the discretion at issue is an equitable discretion of an ad hoc nature similar to the sort of discretion exercised, for instance, in English Chancery Courts from the medieval period to around the 19th Century.³⁵ Unfortunately, this argument betrays a confused understanding of the nature of equitable discretion and its aim to generally do justice. The ad hoc and ‘roguish’ nature of this sort of discretion famously drew the ire of Lord Seldon:

[E]quity is according to the conscience of him that is Chancellor, and as that is larger or narrower, so is equity. ’Tis all one as if they should make the standard for the measure we call a foot, a Chancellor’s foot; what an uncertain measure would this be? One Chancellor has a long foot, another a short foot, a third an indifferent foot: ’tis the same thing in a Chancellor’s conscience.’

Lord Seldon here is lamenting the treatment of similar cases according to vague and arbitrary criteria. Now, when one considers the injunction to do justice, it is difficult to imagine that it would be fair to treat similar cases differently according to undetermined criteria.³⁶ Even when we pursue ‘substantive’ equality and do not presume certain persons equal because of their economic or social position,

³⁰ C Sunstein ‘Beyond Judicial Minimalism’ (2007) 43 *Tulsa Law Review* 825, 829. See also L Tribe *American Constitutional Law* (2nd Edition, 1988) at 1427 and R Bork *Tempting of America: The Political Seduction of the Law* (1990) at 99 and 169–171.

³¹ *National Coalition for Gay and Lesbian Equality and Another v Minister of Justice and Others* [1998] ZACC 15, 1999 (1) SA 6, 1998 (12) BCLR 1517 (CC) or *Lawrence v Texas* 539 US 558 (2003).

³² *Minister of Home Affairs and Another v Fourie and Another* [2005] ZACC 19, 2006 (1) SA 524 (CC), 2006 (3) BCLR 355 (CC); or *Obergefell v Hodges* 576 US __ (2015).

³³ Sunstein (note 30 above) at 825–826. Sunstein goes on in his paper to explain not only the costs of minimalist decision-making but also the instances when this is unjustifiable. He is particularly adamant about instances when minimalists can be called ‘philistines or ostriches’. See Sunstein at 836–841.

³⁴ This will be explored further under the heading of a ‘matter of fact’.

³⁵ J Martin *Hanbury & Martin: Modern Equity* (19th Edition, 2012) 5–15.

³⁶ We need go no further than the notion to treat like cases alike as found in Aristotle’s *Nicomachean Ethics* V.3. 1131a10–b15.

we would still apply general criteria to achieve our conception of social justice.³⁷ Therefore, we cannot treat a case as ad hoc or see it as a single instance where general principles apply in a certain manner only to that decision. This confuses a *particular* thing with an *itemised* thing.

To understand this confusion, one must understand the difference between particular and general things. More to the point, one must understand exactly what a particular thing is. For such aid, we can turn to GEL Owen's interpretation of Aristotle's *Categories*.³⁸ Owen demonstrates how 'particular' things are *wholly determinate* specimens of their class.³⁹ This means that they cannot predicate other things. He argues that to predicate a thing is to classify it or bring it under some more general description.⁴⁰ This means for a thing to be particular, it can no longer be descriptive of further sub-categories of the same kind of thing. Put plainly, we reach a conceptual bottom and cannot go any further when we hit on a particular thing. This is precisely the case when a court applies a low-level principle because of the particular facts and circumstances. It means that we have exhausted our operative facts and defined a case in the narrowest terms possible. But it does not deny that a similar case that is particular in the same way must not be treated the same.

To do so would confuse an item with a particular thing. An item is *an instance* of a particular thing.⁴¹ That is, it is an example of the lowest level of generality possible for the class of the thing under discussion. In moral and legal terms this is the case that the court is currently hearing with those specific plaintiffs before them. So if the narrowest protection we can possibly find for same-sex relationships is decriminalising sexual relations, itemised decision-making would be to treat one case of sodomy differently from another. By parity of reason, there is no reason why even the narrowest instances of deciding on a particular court process in terms of s 173 would call for itemised decision-making. More to the point, one wonders why this would be in the interests of justice. Therefore the *Chief Lesapo* Court was exercising a measure of judicial minimalism by narrowing the application of principle via distinguishing one case from another. It was certainly not arguing that cases ought to be decided according to itemised decision-making. Even the courts of Chancery abandoned this approach.⁴² And

³⁷ C Albertyn 'Substantive Equality and Transformation' (2007) 23 *South African Journal on Human Rights* 255. See especially the discussion on principles and values at 260. I read Prof Albertyn to be arguing for a more comprehensive understanding of inequality in order to determine or achieve equality in legal decision-making. I do not read her denying general values and purposes being applicable to such cases.

³⁸ Specifically 12a17–1.

³⁹ GEL Owen 'Inherence' (1965) *Phronesis* 98.

⁴⁰ Consider the terms 'colour', 'pink' and 'salmon'. By Owen's lights, salmon is the particular thing, whereas pink (mid-level generality) and colour (higher level generality) are general things that predicate the colour or shade of pink called salmon.

⁴¹ Owen (note 39 above) at 104.

⁴² Consider the various general principles of equity like 'he who comes to equity must come with clean hands' (PH Pettit *Equity and the Law of Trusts* (1990) 416). Lord Eldon after his labours as Lord Chancellor said: 'Nothing would inflict on me greater pain in quitting this place than the recollection that I had done anything to justify the reproach that the equity of this court varies like the Chancellor's foot.'

yet the Constitutional Court is at the very least contradictory. In *Mukaddam* the Court seems to reject itemised decision-making by stating:

This does not mean that the court of first instance is not obliged to follow the law as developed by a superior court. In deciding whether a class action should be allowed, that court is bound to apply the standard or test laid down by a superior court. This accords with the principle of judicial precedent. This means that in future the High Courts will be bound to apply the interests of justice standard and in determining where those interests lie in a given case, guidance will be sought from the factors [for certification] mentioned ... above.⁴³

However, in the very application of these factors, the Court states:

These requirements must serve as factors to be taken into account in determining where the interests of justice lie in a particular case. They must not be treated as conditions precedent or jurisdictional facts which must be present before an application for certification may succeed. The absence of one or another requirement must not oblige a court to refuse certification where the interests of justice demand otherwise.⁴⁴

So, there is apparently a standard set, but a court can pick and choose which elements of the standard apply. There are two problems with this line of reasoning. The first is the court has not defined in any meaningful sense what the interests of justice are. It seems whatever form of justice is being discussed is assumed rather than reasoned. The second is if the application of the standard is as widely discretionary as this, notionally a court confronted with a case that is for all intents and purposes analogous to a previous decision need only invoke the undefined interests of justice and claim the *particular* facts demand this. Instead of course, we now know that one is treating these facts as a particular thing at all — but as an item. The Court's standard is illusory and itemised precisely because of its glib treatment of the interests of justice standard and the a la carte approach to the 'factors' for class certification.

If there is to be flexibility in our decision-making, perhaps it would do to be inspired by how Chancery Courts later devised entirely new remedies when confronted with cases where there was a moral wrong done but no legal remedy. Recall that the law of trusts as we know it is entirely a creation of equity courts.⁴⁵ As is the remedy of specific performance.⁴⁶ Saliently, so is the class action.⁴⁷ Thus, itemised and unique decision-making must be rejected.

D Mediating Guidelines are Necessary for Ensuring Distributive Justice

A final objection concerning the facts and circumstances of a case may relate to the Court drawing on a term borrowed from general legal parlance called a

⁴³ *Mukaddam* (note 3 above) at para 47.

⁴⁴ *Ibid* at para 35.

⁴⁵ Martin (note 35 above) at 11.

⁴⁶ Grounded on the maxim that 'equity imputes an intention to fulfil an obligation'. See *ibid* at 32.

⁴⁷ See Story J's remarkable judgment in *West v Randall* (29 F. Cas. 718 (R.I. 1820)) wherein he says: It is a general rule in equity that all persons materially interested, either as plaintiffs or defendants in the subject matter of the bill ought to be made parties to the suit, however numerous they may be.

‘matter of fact’ or a ‘decision on the facts’. That is, when deciding its own process, a court treats this decision as not a matter of law. I shall discuss this concept as it applies in delict, although similar logic may apply beyond this to areas like administrative law review.⁴⁸ In our law of delict, it is now trite that when a court applies the test for negligence, this will depend on the facts and circumstances of a case.⁴⁹ Again, familiar nebulous ‘facts and circumstances’ language is used. As Anton Fagan points out, this passage means that a court has set out a kind of rule that is not developed every time it is applied to a set of facts.⁵⁰ In other words, we can say it grants the court a perpetual discretion to apply a legal rule directly instead of being bound by ‘mediating’ rules created by previous courts.⁵¹ Surely this supports the ‘uniqueness’ argument since a court must decide a case according to criteria that do apply beyond the case before it? It does not. What this does do is simply deny the creation of precedent.⁵² Therefore, no ‘mediating’ *legal* rule is created. This makes it a matter of fact.

However, there can be two types of ‘matters of fact’. There can be a matter of fact that still requires a court to exercise a moral judgement, and a matter of fact that simply does not.⁵³ The test for negligence is the former kind of ‘matter of fact’. It is in essence a rule that requires courts to exercise moral judgements, but does not have the effect of developing subsequent legal rules. This is an instance of the law ‘passing the buck’ from a decision made relying on legal reasoning, based on legal authorities, to moral judgements.⁵⁴ This has the result that a ruling arrived at ‘on the facts’ is to that extent not subject to *legal* generalisation.⁵⁵ However, this does not mean that the decision cannot be morally generalised. Let us not be confused by whom the buck is passed to. In such an instance, a judge is acting not in his capacity as a judicial officer, but as a moral reasoner or ‘finder of fact’.⁵⁶

Because no legal precedent is created, there is accordingly no legal *rule* that is developed when judges decide certain cases ‘on the facts’. It is necessary to make a distinction here between a rule and the reasons or purpose for the rule’s existence. Consider the famous example from the Hart/Fuller debate of a rule prohibiting vehicles being driven in a public park – the purpose of which would be to prevent a park ranger from having to determine using his all-things-considered judgment,⁵⁷ which kinds of visitors she will allow to enter the park.⁵⁸ Similarly, the purpose of precedent is to establish a rule that lower and later courts

⁴⁸ T Endicott ‘Matters of Law’ (1998) 114 *Law Quarterly Review* 292.

⁴⁹ *Kruger v Coetzee* 1966 (2) SA 428 (A) at 428E.

⁵⁰ A Fagan ‘The Secondary Purpose of the Spirit, Purport and Objects of the Bill of Rights’ (2010) 127 *South African Law Journal* 611, 613.

⁵¹ *Ibid.*

⁵² A Fagan ‘Confusions of K’ (2009) 126 *South African Law Journal* 154, 176.

⁵³ *Ibid.* at 176–177.

⁵⁴ J Gardner ‘The Many Faces of the Reasonable Person’ (2015) 131 *Law Quarterly Review* 563, 568.

⁵⁵ *Ibid.* at 569.

⁵⁶ *Ibid.* at 570.

⁵⁷ All-things-considered including the moral, social and practical reasons for the rule. See F Schauer *Playing by the Rules: A Philosophical Examination of Rule-Based Decision-Making in Law and in Life* (1991) 77–79.

⁵⁸ See HLA Hart ‘The Separation between Law and Morals’ (1958) 71 *Harvard Law Review* 593, 603.

would, ordinarily,⁵⁹ apply without exercising a judgement as to the soundness of the rule itself.⁶⁰ When courts decide cases ‘on the facts’ they mean to allow lower (and later) courts the discretion to exercise their all-things-considered judgement, taking into account the purpose, social and especially moral reasons for having a general rule like ‘regulating court process in the interest of justice’. However, this does not mean that when these later courts have to make decisions they are left at sea.

Usually when courts establish a precedent from a general rule, they are refining that general rule to suit a particular sub-category of cases.⁶¹ Here the notion of the general rule controls part of the formulation of the subsequent mediating rule, but does not wholly determine it.⁶² The decision-maker has to exercise a discretion from a range of options and settle on one refinement.⁶³ Yet, we have just stated that decisions ‘on the facts’ do not create these refined mediating rules. This remains true. However, there are subsequent mediating guidelines, not legal rules, which assist courts in making their all-things-considered judgements. I shall call these mediating moral judgements. These mediating moral judgements are a form of moral precedent that act purely as instruments that give logical guidance to subsequent courts, although they are not rules.

A case in point from the law of delict is how the test for negligence has been refined by mediating moral judgements. Consider the original formulation of negligence in our law by Innes CJ in *Paine*’s⁶⁴ case:

Once it is established the danger would have been foreseen and guarded against by the diligens paterfamilias, the duty to take care is established and it remains only to determine if it has been discharged.⁶⁵

Now consider how Shreiner JA then elaborated on this test some 30 years later:⁶⁶

⁵⁹ I say ‘ordinarily’ to make exception for when a rule in our system is constitutionally challenged or if it is a customary or common-law rule which is developed. I also make allowance for when a rule is distinguished and so its scope of application is narrowed.

⁶⁰ A legal rule, in the form of precedent, would give what Raz has in a different context called a ‘protected’ reason for subsequent courts to follow. This would be a reason not only to act in a certain way, but also to exclude competing reasons for alternative action. This exclusionary effect is exemplified by how difficult it is for courts to directly overturn previous precedent or for them to go through the difficult enterprise of ‘distinguishing’ a case, thereby adding a plausible condition to an existing legal rule. The force of precedent is such that even where the Constitutional Court does depart from previous decisions it often is loath to say so outright. See the discussion of the *Gcaba v Minister for Safety and Security and Others* [2009] ZACC 26, 2010 (1) SA 238 (CC), 2010 (1) BCLR 35 (CC) and *Chirwa v Transnet Ltd and Others* [2007] ZACC 23, 2008 (4) SA 367 (CC), 2008 (3) BCLR 251 (CC) in J Brickhill ‘Precedent and the Constitutional Court’ (2010) 3 *Constitutional Court Review* 81, 81–86. Cameron JA has expressed concern over this habit by the Court in *True Motives 84 (Pty) Ltd v Madhi and Another* [2009] ZASCA 4, 2009 (4) SA 153 (SCA), 2009 (7) BCLR 712 (SCA) at para 102 fns 52 and 53.

⁶¹ See J Finnis’ example of *determinatio* (implementations of general directives) in his *Natural Law and Natural Rights* (1980) 284–285. These are the sorts of rules that require a decision-maker to exercise a discretion in setting them down from a more general rule.

⁶² Ibid.

⁶³ For example, in the case of our no vehicles rule, deciding affirmatively that a skateboard is a vehicle is an instance of *determinatio*.

⁶⁴ *Cape Town Municipality v Paine* 1923 AD 207.

⁶⁵ Ibid at 217.

⁶⁶ *Herschel v Mrupe* 1954 (3) SA 464 at 477A–477B.

Apart from the cost or difficulty of taking precautions, which may be a factor to be considered by the reasonable man, there are two variables, the seriousness of the harm and the chances of its happening. If the harm would probably be serious if it happened the reasonable man would guard against it unless the chances of its happening were very slight.

Note that the Court has now added further considerations as to how to determine the reasonableness of preventative steps taken. Shreiner JA relies heavily on previous decisions whilst grappling with how to determine preventability and foreseeability.⁶⁷ The SCA then crystallised these considerations for testing foreseeability and preventability into four factors that inform this assessment.⁶⁸ This process of exposition and refinement is done *despite* the test for negligence being a moral judgement made by ‘balancing’ various competing moral considerations.⁶⁹ Therefore, although aware that these previous decisions are not precedent and binding, courts still refer to them and do not simply ignore these previous judgments.⁷⁰ Courts do not simply assume that in order to broadly determine negligence they can pick in an arbitrary manner which factors they would like to apply. There remains a need to justify their reasoning. More importantly, courts do so with reference to *previous decisions* despite this being a matter of fact, involving all-things-considered moral judgements. This peculiar reliance on previous decisions in spite of their being ‘non-binding’ precedent is what I mean by mediating moral judgements. There remains a process by which the common law ‘works itself pure’,⁷¹ even where there is no subsequent law being created.

We see a similar trend in how our Constitutional Court has interpreted s 173 of the Constitution. For instance, we have a legal rule stating ‘court rules ought to be applied following the interests of justice’.⁷² But, when the Constitutional Court is faced with a case where a litigant had her case dismissed in the SCA because she flouted the SCA’s procedural rules, there would be a mediating guideline to the effect that ‘it is not in the interests of justice to grant leave to appeal where the case was dismissed in the court a quo because a litigant grossly flouted that court’s rules without good reason’.⁷³ Thus, when the Court is faced

⁶⁷ Ibid at 475B–475C.

⁶⁸ *Ngubane v South African Transport Services* [1990] ZASCA 148, 1991 (1) SA 756 (A) at 776H–J and *Cape Metropolitan Council v Graham* [2000] ZASCA 93, 2001 (1) SA 1197 (SCA) at 1203G–H.

⁶⁹ Ibid. Further, the use of the word ‘balancing’ is often misleading. It has undertones of particularised decision-making and some form of ‘uniqueness’ argument. When ‘balancing’ moral considerations a court is still deciding that when faced with certain circumstances one moral consideration is more compelling than another. We cannot wish away the general nature of our moral reasoning.

⁷⁰ See how the court in *Ngubane* (note 68 above) expressly argues that two considerations used to decide preventability are not relevant because they are met. The court does not simply assume because this is a value judgement it can simply pick for itself which factors to apply – the court still feels the need to justify not applying these factors.

⁷¹ The famous phrase by Lord Mansfield in *Omychund v Barker* 26 (1744) ER 15 at 23.

⁷² Section 173 of the Constitution.

⁷³ This is indeed O’Regan J’s holding in *Mabaso v Law Society, Northern Provinces, and Another* [2004] ZACC 8, 2005 (2) SA 117 (CC), 2005 (2) BCLR 129 (CC) and re-affirmed by Froneman J writing separately in *Dengetenge Holdings (Pty) Ltd v Southern Sphere Mining and Development Co Ltd and Others* [2013] ZACC 48, 2014 (3) BCLR 265 (CC), 2014 (5) SA 138 (CC) at para 143.

with a case where a litigant failed to apply for condonation from the SCA in time without good reason, they would apply the mediating guideline and by so doing serve the interests of justice. So why do courts feel the need to refer to previous decisions even when these are not precedent? Why continue looking backwards for guidance?

There are two plausible and related explanations as to why courts still look backwards at previous decisions even though they are not bound by precedent. The first is an explanation rooted in the nature of reasoning in general – namely, that it is entirely *rational* for courts to look backwards. We ought to understand that principles of rationality in practical reasoning are instrumental.⁷⁴ That is, these are principles we ought to adopt if they aid us in better securing whatever ends we happen to have, putting to one side the intrinsic value of those ends.⁷⁵ Therefore, if we assume that the ‘ends’ judges aim for are to make correct moral decisions on what sort of conduct a reasonable person ought to foresee and take steps to avoid, previous decisions are useful instruments in securing these ends.⁷⁶

The second explanation as to why judges look backwards is because the rule of law demands this. Even though the law has passed the buck to the purely moral sphere, judges are still concerned that no legal vacuum exists that allows for the arbitrary exercise of public power.⁷⁷ Importantly, this does not make this a legal decision. The law is only concerned with vesting this further moral reasoning with the normative value of the rule of law. It may seem bizarre that the rule of law is seen as a normative value – but it most certainly is. Gardner has shown that it does not follow that for something to be a law it must necessarily comply with the rule of law in order to qualify as a law.⁷⁸ Notionally, we can imagine a legal system that does not guide its population by ensuring the predictability of government actions.⁷⁹ In fact, South Africa could easily have such a legal system that creates rules only for a few officials, who in turn arbitrarily decide the further actions of the populace on an ad hoc basis. The law would then goad subjects into compliance through an economy of threats, rather than guiding their actions.⁸⁰ The court, by creating this pocket of moral reasoning, has perhaps come close to doing so. However, aware that the central case of law is its complying with the rule of law,⁸¹ courts are anxious to re-introduce the rule of law even where there is no law.⁸² Hence, we have our courts looking backwards at previous judgments to ensure some measure of predictability as to what a reasonable person may do

⁷⁴ Parfit (note 19 above) 1984 at 4.

⁷⁵ John Gardner ‘Justifications and Excuses’ in his *Offences and Defences* (2007) at 101 fn19.

⁷⁶ The rationality of these guidelines is also premised on the fact that even moral reasoning is seldom particularised. See J Raz ‘Review: The Trouble with Particularism (Dancy’s Version)’ (2006) 115 *Mind* 99–120.

⁷⁷ This is likely informed by the fear that this non-rule based reasoning would swallow the entirety of a legal system, best expressed by Hart’s example of a ‘scorer’s discretion’ in his *Concept of Law* (note 20 above) at 138–141.

⁷⁸ J Gardner ‘Hart on Legality, Justice and Morality’ in his *Law as a Leap of Faith: Essays on the Law in General* (2012) 8.

⁷⁹ Ibid.

⁸⁰ HLA Hart *Punishment and Responsibility* (1978) 40–44.

⁸¹ Gardner (note 78 above) at 8.

⁸² Gardner (note 54 above) at 582–584.

when faced with the prospect of causing harm to others. The same applies when courts have to make a moral judgement as to whether it serves our notions of justice to allow certain cases to be heard by a court, considering limited judicial resources.

The value of such predictability required by the s 173 power has a surprising advocate – Justice Albie Sachs. In his separate opinion in *South African Broadcasting Corporation v National Prosecuting Authority*,⁸³ he compellingly argues:

The reconciliation of all the different interests involved cannot be achieved by privileging one interest over another. *Nor can it be accomplished by leaving each case to be determined in an ad hoc manner according to the robustness or sensitivity of the Judges concerned.* ... Clear guidelines need to be established in advance so as to provide a principled and functionally operational basis for the granting or refusal of access to the electronic media ... As I see it, such guidelines may well give to courts a certain margin of appreciation in terms of the application of the guidelines on a case-by-case basis. *Pre-established and principled guidelines, subject to periodic review, would assist broadcasters in their planning.*⁸⁴

Here Sachs J not only argues that guidelines created by the exercise of the s 173 power ought to be principled, presumably grounded in rational decision-making, but also that these guidelines ought to be clear and not ad hoc in order to allow citizens to plan their activities. It is now somewhat trite that the rule of law is generally intended to allow individuals to plan their lives, in our case their litigation, in order to avoid capricious state decision-making. Thus, mediating moral judgements have mostly instrumental functions.⁸⁵ First, they are the sorts of judgments that have proven effective at allowing judges to determine correct moral decisions concerning reasonable moral behaviour (in the context of negligence) or deciding on just distributions of legal resources. Second, they aim to give guidance to subsequent courts, and in turn the general public, as to what the state will regard as reasonable harm-causing action and just procedural distribution of legal resources. They assist in aiding citizens to plan their lives according to mostly predictable criteria upon which to anticipate state (especially court) decisions. These instrumental principles, which for our purposes do not have to be compelling ends themselves, allow the public (or at least their lawyers) to better predict what will count as reasonable harm-causing events or what judges will consider just process in litigation.⁸⁶ However, this only explains why these guidelines are useful and perhaps necessary to realise the rule of law. We must still explain why we need to apply these mediating guidelines in pursuing the interests of justice.

⁸³ *South African Broadcasting Corporation Ltd v National Director of Public Prosecutions and Others* [2006] ZACC 15, 2007 (1) SA 523 (CC), 2007 (2) BCLR 167 (CC).

⁸⁴ *Ibid* at para 151 (emphasis added).

⁸⁵ I say ‘mostly’ because Nozick has shown that rational principles can also acquire their own inherent or intrinsic value. See R Nozick *The Nature of Rationality* (1993) at 133–139.

⁸⁶ I accept Nozick’s refrain that rationality does not necessarily have exclusively instrumental value. It is plausible that over time the mediating guideline of considering the utility of an activity in assessing whether a reasonable person would take steps to avoid harm may acquire symbolic value independent of its usefulness in helping a judge determine the actions of a reasonable person. This guideline may become an indicator that a judge who expressly engages with this guideline in detail is an especially conscientious and morally astute decision-maker. See Nozick (*ibid*) on the symbolic value of rationality.

E The Meaning of 'Justice' in Section 173

To understand the link between mediating guidelines and the interests of justice, we need an understanding of what the Constitution means by the word 'justice' in s 173. Now, this may seem rather obvious, yet some lawyers may be hesitant to give meaning to seemingly uncontroversial words like 'when' in constitutional text.⁸⁷ This may be even more controversial when we are dealing with concepts as loaded and contested as 'justice'. But let us strain ourselves. There is a deep and troubling problem with an 'interest of justice' standard applied haphazardly by judges who have no specific notion of exactly what 'justice' requires beyond inexplicable impulse. We have seen the emptiness of overstated platitudes about 'facts and circumstances' and vague gestures towards deciding things 'on the facts'. Even if one disagrees with the arguments against itemised decision-making, we must still ask on what reasons the Constitutional Court relies when castigating the SCA and all other courts that 'rigidity has no place in our law'. This must be because the Court believes its 'flexible' standard is demanded by the interests of justice. It follows, therefore, that the Court *already* holds some notion of justice it wishes lower courts to consider. This alone warrants us exploring what *form* of justice is demanded by our Constitution.

I submit that the form of justice demanded by s 173 includes, but is not limited to, distributive justice. I say 'includes, but is not limited to' because there are other forms of justice.⁸⁸ This includes what some call 'corrective'⁸⁹ and others 'commutative'⁹⁰ justice. Whichever term one prefers, the point is to stress that corrective or commutative justice concerns justice as done exclusively between the parties.⁹¹ We are concerned, however, with the form of justice that concerns how common property, namely court resources, are distributed. We focus, therefore, on a distributive claim.

At the root of the flexibility standard is a genuine concern about the right of access to courts.⁹² The logic is that to not allow claims to be heard in a court risks encouraging claimants to seek self-help and violate the rule of law.⁹³ This is all true. However, this observation also reveals something about what the interests of justice demands, which is more complicated than an imprimatur to 'allow as many claims as possible lest mayhem reign'. Mayhem is only an option because there are moral rights still at stake when individuals bring claims before a court, to which people remain attached enough to take the law into their own hands should they not be heard. They believe they have a *moral* right to be compensated

⁸⁷ See D Davis 'How Many Positivist Legal Philosophers Can be Made to Dance on the Head of a Pin? A Reply to Professor Fagan' (2012) 129 *South African Law Journal* 59.

⁸⁸ Aristotle *Nicomachean Ethics* 5; 2; 30 (Thomas, Trendick and Barnes translation, Penguin Books 2004) at 118. See also Thomas Aquinas *Summa Theologiae* 37; ques. 61; art. 1 in the Reply at 3 for the passage on common goods.

⁸⁹ J Gardner 'Finnis on Justice' in J Keown & RP George (eds) *Reason, Morality, and Law: The Philosophy of John Finnis* (2013) 171.

⁹⁰ Thomas Aquinas *Summa Theologiae* 37; ques 61; art 1. This term of art was explained further by Finnis in his *Natural Law and Natural Rights* (note 61 above) at 177.

⁹¹ Finnis (ibid).

⁹² Section 34 of the Constitution.

⁹³ *Chief Lesapo* (note 1 above) at para 11. Restated and approved in *Mukaddam* (note 3 above) at para 30.

for some harm caused, promise unfulfilled, or administrative decision that adversely impacts on an existing right or legitimate expectation. This is no mean point. For people to feel that they have a right to flout the rule of law, they presumably believe that they are morally justified to do so. Therefore, there is a natural tension between the rule of law and moral claims that procedural law aims to reconcile. This is done by institutionalising or recognising these moral claims as legal claims that can be decided by a court.⁹⁴ Whether institutionalisation of these moral rights has been done well or justly is a matter of distributive justice.⁹⁵ Gardner has shown that institutionalisation occurs when the law recognises these moral claims are worthy of legal protection.⁹⁶ He also astutely shows that there is a *second* distributive decision as to whether the law ought to recognise a certain moral claim as a *type* of cause of action, ie as a delictual or contractual right.⁹⁷

I argue that there is a *third* distributive decision that must be taken when courts have to decide whether certain claimants, who have a cause of action, may avail themselves of certain procedural mechanisms such as class actions. This sort of decision is aimed at matching certain causes of action with certain procedural instruments. The question here is not whether a litigant has a justiciable right per se, but rather if this right may be vindicated using this or that mechanism. Therefore, a court must exercise a judgment. But by what criteria may a court exercise this sort of judgment?

Comprehend the nature of this distributive decision. It is a decision as to which sorts of claims are best facilitated by a certain procedural instrument. Now ordinarily we leave this sort of decision to a litigant. This may be premised on the notion that a litigant is best placed to decide for themselves the best strategy to vindicate their right to compensation or performance. So then, why should we require further guidance from courts?

F Perfectionism and the Importance of Mediating Guidelines

The answer to the above question lies in a perfectionist view of distributive justice. A perfectionist form of distributive justice aims to support, rather than supplant, the decisions citizens make.⁹⁸ The state aims to assist its citizens by promoting their welfare.⁹⁹ If we assume that attaining appropriate redress promotes a citizen's welfare, then it is perfectly legitimate for a court to assist such a claimant in choosing the best procedural mechanism to vindicate their right. However, a court cannot make this choice for a claimant and so instead it may opt to assist a claimant in two ways. The first is by providing the claimant with sufficient guidance that she may generally predict that if she elects to bring her claim in the

⁹⁴ J Gardner 'What is Tort Law for? Part 2. The Place of Distributive Justice' *Oxford Legal Studies Research Paper No. 62/2013* (24 May 2013) at 8, available at <https://ssrn.com/abstract=2269615> or <http://dx.doi.org/10.2139/ssrn.2269615> at 8.

⁹⁵ *Ibid* at 8–9.

⁹⁶ *Ibid*.

⁹⁷ *Ibid* at 10.

⁹⁸ J Oberdiek 'Perfecting Distributive Justice' (25 September 2014) 5, available at <https://ssrn.com/abstract=2501444> or <http://dx.doi.org/10.2139/ssrn.2501444>. See also J Raz 'Liberty and Trust' in RP George (ed) *Natural Law, Liberalism, and Morality* (1996) 113.

⁹⁹ Raz (note 98 above).

form of a class action, or take her case on appeal, then her chances of success are good or bad. The role of our mediating guidelines is to assist claimants in this respect.

If we recall that the purpose or ends of the s 173 power is to allow courts' discretion to be exercised in the interests of justice, then we see how these mediating guidelines are not only optional, but necessary. Instrumental principles are constitutive of the ends they pursue.¹⁰⁰ This specific aspect of instrumental reason is often misunderstood. We may be tempted to think that we first pick our ends, decide whether it is worth it to realise these ends and then choose the means to realise our ends. However, to think this way would miss that to will an end is also to will to realise that end.¹⁰¹ This means in a lot of cases when we pick our ends we also pick our means. By giving guidance as to which cases are best brought as class actions or taken on appeal, these mediating guidelines are the sort of instrumental principles that are pre-supposed by choosing justice as the ends of the s 173 power. That is, these mediating judgments are constitutive of justice by being efficient at doing justice.¹⁰² Importantly, these instrumental mediating guidelines should not only be efficient at doing corrective justice, allowing those who have suffered harm to claim; they should also be efficient at doing distributive justice. Such efficiency at doing distributive justice is achieved by creating sufficient predictability to allow for proper planning of litigation and to pass on the initial distributive decision to plaintiffs at the stage when they decide which procedural instrument to choose. But this assumes that the criteria themselves are sound distributive instruments.

This brings us to the second manner in which courts promote the welfare of litigants. Namely, by tailoring certain procedural requirements to suit certain classes of litigants. In this way, a court subsidises certain claims it deems worthy of being given priority amongst others. However, in order to do so, we must keep in mind the nature and purpose of the procedural instruments in question. Let us consider two categories of possible claimants who want to bring their claims as class actions. Category A are claimants who have suffered harm caused by the same event and cannot afford to litigate their case. Category B are claimants who could notionally afford their litigation, but do not want to bear the administrative costs of organising potential claimants.

Now recall our case of bread price-fixing. If we accept that the purpose of a class action is to allow multiple claims with similar questions of fact and law to be answered in one swoop,¹⁰³ we can begin to determine which of the above categories deserve certification. Remember that there are opt-out class actions available. These are involuntary. By virtue of determining the class, we make individuals a party to the litigation. This is how Category A's lack of financial

¹⁰⁰ C Korsgaard 'The Normativity of Instrumental Reason' in G Cullity & B Gaut (eds) *Ethics and Practical Reason* (1997) 244–245.

¹⁰¹ Ibid.

¹⁰² On constituting justice through instrumental means see J Gardner 'What is Tort Law For? Part 1' (2011) 30 *Law and Philosophy* 1, 18–20.

¹⁰³ *Mukaddam* (note 3 above) at para 10.

means is included in the mediating guidelines aimed at efficiently doing justice. In fact the SCA has recognised as much in *Ngxuzza*.¹⁰⁴ Cameron JA pointedly says:

The law is a scarce resource in South Africa. This case shows that justice is even harder to come by. It concerns the ways in which the poorest in our country are to be permitted access to both ... The procedure has particular utility where a large group of plaintiffs each has a small claim that may be difficult or impossible to pursue individually.¹⁰⁵

Notice how the general mediating guideline of common facts and law is met by virtue of this being a class action. However, our concern for claimants who cannot afford litigation is met by the *type* of class action – an opt-out action. Those who cannot afford litigation are automatically made parties to the litigation at no expense to them. Therefore, the consumers in our bread price-fixing scandal easily include those consumers who could not have brought claims on their own accord.

The involuntary nature of opt-out class actions speak directly to a lack of financial means not only to pursue worthwhile litigation, but also to cover the administrative cost of issuing notices to potential class members in order for them to withdraw if they so wish. But what of poor Mr Mukaddam and his fellow distributors should they be strapped for cash? The simple answer is that there is nothing barring them from bringing such an opt-out class action.

This then poses the question as to why we should allow opt-in class actions. Like joinders, these actions are voluntarily constructed by litigants also in an attempt to decide multiple, similar claims in a single swoop.¹⁰⁶ The only difference between opt-in class actions and joinders is that the former action would immunise plaintiffs against personal liability for costs.¹⁰⁷ Interestingly, Nugent JA in *Mukaddam* before the SCA ruled that opt-in class actions ought to be granted only in exceptional circumstances since we already have joinders available.¹⁰⁸ The Constitutional Court, expressing some bafflement, then took umbrage to this barrier to opt-in class actions.¹⁰⁹ The Court was misguided to do so. Recall our two categories of litigants. Category B is clearly the class in question here. These are litigants who are not so hard pressed financially that they do not need to seek refuge in opt-out class actions. Perhaps distributors such as Mr Mukaddam do not wish to face administrative costs for advertising class action

¹⁰⁴ *Permanent Secretary Department of Welfare, Eastern Cape Provincial Government and Another v Ngxuzza and Others* [2001] ZASCA 85 at para 1.

¹⁰⁵ *Ibid* at paras 1 and 5.

¹⁰⁶ Rule 10(1) of the Uniform Court Rules defines a joinder as:

‘Any number of persons, each of whom has a claim, whether jointly, jointly and severally, separately or in the alternative, may join as plaintiffs in one action against the same defendant or defendants against whom any one or more of such persons proposing to join as plaintiffs would, if he brought a separate action, be entitled to bring such action, provided that the right to relief of the persons proposing to join as plaintiffs depends upon the determination of substantially the same question of law or fact which, if separate actions were instituted, would arise on each action, and provided that there may be a joinder conditionally upon the claim of any other plaintiff failing.’

For opt-in class actions see *Mukaddam* (note 3 above) at para 23.

¹⁰⁷ *Mukaddam* SCA (note 6 above) at para 14.

¹⁰⁸ *Ibid* at para 12.

¹⁰⁹ *Mukaddam* (note 3 above) at para 51–55.

notices in newspapers, television and other mediums. Instead they would rather respondents carry these costs. Now whatever the merits of their corrective justice claims, there is an obvious perversity in forcing respondents who are being sued to be obliged to publically beg to be sued. Yet, this is not an objection founded in distributive justice. The distributive concern is best expressed by Nugent JA:

The potential for personal liability for costs will often serve as a salutary restraint upon frivolous actions that are brought oppressively for the purpose of inducing defendants into financial settlements, which is one of the dangers to be avoided in certifying class actions.¹¹⁰

Here the Court is concerned about preventing the abuse of a common good.¹¹¹ This common good is a precious and scarce court resource. But is it fair to have victims carry the cost of notification even if such victims can afford it? Nugent JA parries this attack by pointing out:

Indeed, the court that becomes seized of the case has a wide discretion to determine where the costs should fall, taking account the merit of the claim and the conduct of the litigation, and is better placed to do so than a certifying court.¹¹²

A court that hears these merits is better placed precisely because at the end of the litigation, after the merits have been decided and their worth tested, that court will have determined whether the claimants are in fact victims worthy of such benefits. Therefore, the state supports the welfare of litigants and ensures sound norms of distributive justice by attaching a load to an applicant's back in their uphill battle, whilst offering them reward in the form of eventual recompense for all costs, including advertising costs, should they litigate in good faith.

Is it a bad thing to have as many procedural options as possible? Perhaps the more procedural mechanisms we have available, the better. This is false. To do so misses what certification requirements are. They are mediating guidelines aimed at doing justice. This of course includes corrective or commutative justice, embodied by the original cause of action. With regards to class actions, mediating guidelines achieve corrective or commutative justice by requiring a litigant to disclose a cause of action.¹¹³ But mediating guidelines in the form of certification criteria also aim at achieving distributive goals. A court has a justifiable interest in ensuring that a common good like court resources ought not to be abused. Recall that a government ought to only subsidise worthy choices or good decision-making. Allowing court resources to be co-opted in pursuit of abusive objects is not a legitimate function of government. These concerns do

¹¹⁰ *Mukaddam SCA* (note 6 above) at para 14.

¹¹¹ Finnis *Natural Law and Natural Rights* (note 61 above) at 179.

¹¹² *Mukaddam SCA* (note 6 above) at para 14.

¹¹³ *Trustees for the time being of Children's Resource Centre Trust* (note 5 above) at para 34–43. See also Mhlantla AJ's partial dissent in *Mukaddam*:

‘Certification is also significant in protecting the interests of persons whose right to pursue a claim may be extinguished by a class action. The outcome of a class action, favourable or unfavourable, is binding on all members of a class. Thus, the right of those members to raise the dispute again will, in terms of our current law, be substantially limited by the application of the *res judicata* principle.’

not apply to joinder proceedings. Therefore, proponents of the argument that having as many procedural mechanisms as possible is a positive thing have to argue either that there is no appreciable risk of abuse, or that the state must still underwrite or support abusive practices. The former may be controverted by empirical evidence and once this is done, it serves as reason to dispense with the exceptional circumstances barrier. The point remains though that Nugent JA provided a good reason to treat opt-in class actions differently to joinders *prior* to costs being decided. This shifts the burden of justification to his opponents. As for any suggestion that the state ought to underwrite abusive practices, placing opt-in class actions on the same par as joinders increases the chances that many would be drawn to not paying the administrative costs of class notices.

But the Constitutional Court did not agree. They dispensed with the ‘exceptional circumstances’ criterion. And the Court did not stop there. Instead, they gave us what they called the ‘flexible approach’. This approach seems difficult to define, but at a stab it rails against some vague formalist approach of seeing the certification criteria as requirements. Apparently courts ought to see these criteria as ‘factors’ to be applied according to the facts and circumstances of a case in a court’s exercise of the s 173 power. Now there are two broad objections to this view. This is over and above misunderstanding the nature of the s 173 power and what decisions according to facts and circumstance mean when applying this discretion. This we have already addressed and shall not repeat.

IV FALLACY II: FLEXIBILITY CAN CONCEPTUALLY DISPENSE WITH CERTIFICATION REQUIREMENTS

A The Purpose of Certification Requirements for Class Actions

A further objection to this flexible approach is a misunderstanding of the nature of class actions. As we have said, the purpose of a class action is to allow multiple claims to be heard all at once. A complementary purpose is to remove financial litigation burdens from those who cannot afford to litigate. One can disagree with whether a certain criterion has been met — specifically, one can take issue with the evidential standard for the criterion. However, it is another thing entirely to take issue with the criterion itself applying. This begs the question as to which criteria are disposable.

Imagine a group of Kayelitsha residents who are the victims of illegal dumping of toxic waste by several companies in nearby landfills. Exactly how is it against the interests of either distributive or corrective justice to require common issues of fact and law? For opt-out class actions, this is how we define the class and draw in all residents in Kayelitsha as parties to the litigation. And can it be in the interests of justice to not require these applicants disclose a cause of action,¹¹⁴ instead of arriving at the merits trial only to discover some serious legal impediment that would require constitutional challenge of a legal doctrine or right to change? Can we have a class action without a class representative(s) or an unsuitable representative?¹¹⁵ In the absence of some collective hive mind, or a tolerance

¹¹⁴ Ibid.

¹¹⁵ Ibid.

for an opportunistic class representative who acts against the interests of the class, this is doubtful. The answer to all these questions is in the negative not only because the questions were rhetorical. It is because each of the certification criteria are *constitutive* of the very concept of a class action. By this I mean that they are indispensable requirements of a class action because they are instrumental in facilitating multiple claims by different individuals to be heard in a single case. To this end, we *need* common questions of fact and law, disclosing a cause of action, which can be litigated by a representative(s). This is not conceptually optional. Wallis JA in *Children's Resource Centre Trust* gestures towards something like the constitutive nature of these criteria.¹¹⁶

But the 'exceptional circumstances' criteria is not constitutive of the notion of a class action. For that matter neither is the requirement that, given the composition of the class and the nature of the proposed action, a class action is the most appropriate means of determining the claims of class members.¹¹⁷ These criteria both assume that it is possible to construct a class action. However, the focus of the appropriateness criterion is a comparison of class actions *qua* individual litigation. Consider the example Wallis JA gives:

A class action may be certified in respect of limited issues, for example, negligence in a mass personal injuries claim, leaving issues personal to the members of the class, such as damages, to be resolved separately.

The predominant form of justice aimed at by the appropriateness criterion is corrective. This is because our concern for the very personal nature of the damages claim is that some victims may have suffered greater damage than others. Consider for instance our distributors in *Mukaddam* where some distributors may be able to pass on greater parts of the loss incurred by producer's price-fixing than others. However, as previously stated, our exceptional circumstances criterion is concerned with a greater distributive question of the potential abuse of a common good.

A final point on the purpose of class actions warrants some consideration of the purpose of certification itself. Can we have a court shirk its duty to regulate its own process in the interests of justice by discarding the very real need for certification in instances of constitutional damages for Bill of Rights claims? The assumption must be that certification is a hindrance rather than a help in allowing class action claimants to pursue corrective redress for rights violations. Heed then Mhlantla AJ's (as she was then) thoughtful refrain in her partial dissent:

The preliminary stage of certification ... plays an important role in informing and protecting potential class members through, for example, notification procedures. Such notification procedures are not available in joinder proceedings. By contrast, once a class is certified a court must provide for members of the class to be notified of the upcoming action. Notification procedures are particularly significant in the context of opt-out class actions because all members of the class will be bound by the judgment except for those members who actively opt out of the class. Without adequate notification procedures

¹¹⁶ Ibid at para 26: 'There is an element of overlapping in these requirements. For example, the composition of the class cannot be determined without considering the nature of the claim.'

¹¹⁷ Ibid.

individuals could be bound to a judgment even though they had not explicitly consented to, and may not even have been aware of, the action.¹¹⁸

So what is it about Bill of Rights claims that they do not deserve the benefits of certification? Is it some aesthetic revulsion at requiring procedural oversight over complex litigation in the interests of those whose rights have been violated? The Court is silent on this question. Hopefully, this is only prudent judicial minimalism. If so, when the Court finally confronts the question as to whether constitutional damages claims ought to be certified, they must address Mhlantla AJ's correct exposition of the purpose of certification. Again, to stress my earlier point that judicial minimalism is not my quarry in this article, the suggestion is that a court must explain why, appreciating the purpose of certification, claims for constitutional damages do not deserve the benefits of prior review that all other claims receive.

B Confusions on Evidential Standards

The third objection is that the Court confuses 'flexibility' in the standard of compliance of a requirement with flexibility in dispensing with a requirement all together. In class action litigation, the former is quite common, but the latter is almost unheard of. For instance, in the United States there is no disagreement by the Supreme Court on whether all the requirements for a class action must be met. This is despite the sometimes heated divide on controversial class actions in the civil rights arena.

Consider for instance the now famous *Wal-Mart v Dukes*¹¹⁹ decision. This litigation was the largest class action in US history brought by 1.5 million female employees against Wal-Mart stores as a gender discrimination claim concerning pay and promotion matters.¹²⁰ Importantly, there was no evidence of a formal or informal policy of discrimination adopted by Wal-Mart – the charge instead was that there was a nationwide gendered disparate impact.¹²¹ Scalia J, writing for the majority, dismissed the request for certification because there were no common issues of law and fact since local supervisors had discretion as to their promotion and pay decisions.¹²² More to the point, plaintiffs did not show there was any reason to believe there was a common mode of exercising this discretion that targeted female employees, especially since the employees were all at different levels of employment, across varying regions.¹²³ The crux for Scalia J was that there were no common issues that predominated such that the class action could be decided in one stroke.¹²⁴ Ginsberg J, however, argued commonality is met if there are similar results, regardless of the manner in which the discretion is exercised.¹²⁵

¹¹⁸ *Mukaddam* at (note 3 above) at para 61 and fn 44.

¹¹⁹ 564 US (2011).

¹²⁰ *Ibid.*

¹²¹ *Ibid* at 4.

¹²² *Ibid* at 14.

¹²³ *Ibid* at 16.

¹²⁴ *Ibid* at 19.

¹²⁵ *Ibid*, Ginsberg J dissenting in part at 11.

For our discussion, we care only that despite the marked difference in approach, both judgments accepted the commonality requirement as uncontroversial. They simply disagreed on the evidential standard one had to show for such an action in order to prove commonality. If the ‘flexibility’ our Court was referring to was Ginsberg J’s approach, then there would be reason to consider this lowered evidential standard as ensuring the door is not shut to new forms of classes. However, our Court proceeds to gut the very notion of a class action by arguing that necessary criteria like commonality are optional. Recall that this is not an argument for judicial minimalism where a court is deciding the category of cases before it and only that category. Therefore, the sort of ‘flexibility’ we ought to be striving for is precisely the flexibility akin to English Equity courts after the Judicature Acts had been passed. All courts can fashion new procedural remedies when a particularly difficult problem which demands a corrective remedy and the current procedural mechanism do not cover this challenge. We begin to develop mediating guidelines to assist future courts in their allocation of such a new remedy. We do not, however, overhaul existing remedies in so careless a manner that we deny any mediating guidelines exist and so require lower courts to decide cases according to instinctively arbitrary ‘factors’. Further, by focusing on the evidential standards, we include uncovered new cases, whilst still allowing our remedies to be effective at doing not only corrective justice in the case before the Court, but also doing distributive justice at allocating a scarce common good – legal proceedings. This is not anything new to our procedural law. In fact our courts have always had this ‘flexible’ approach before being seized by this haphazard incarnation of ‘flexibility’. Ponder Voet’s observations:

Yet liberty in interpretation is not wholly denied to lawyers or judges; although it lacks the virtue of law. Either they may extend the legal rule to similar cases, its basis [*ratio*] having been considered, that basis operating in other cases tending towards the same utility; even if the law uses restrictive words, since such narrowness and restriction works only to exclude different cases, not those in which the same or similar or greater (stronger) basis [*ratio*] is found.¹²⁶

So, where the requirements by virtue of their words exclude certain cases, a court can fashion a new remedy distinct from a general class action to cater for those plaintiffs. There is no moral justification to take a sceptical wrecking ball in the form of a certification-criteria-is-optional-depending-on-the-circumstances approach to cater for exceptions that can be addressed when confronted by the liberty in interpretation a court possesses by virtue of the s 173 power to act in the interests of justice.

V FALLACY III: FLEXIBILITY EQUALS RIGIDITY

A The Spectre of the ‘Rigid’ Judge

Finally, we come to the last objection concerning the Constitutional Court’s description of the flexibility approach – there is no clear evidence of the ‘rigid’ adversary in our law to rail against. The Court ought to show at what point in our

¹²⁶ J Voet *Commentarius ad Pandectas* 1:3:20. I am indebted to Prof Helen Scott for this translation.

law, when exercising the inherent power of courts to oversee their own process, have courts advocated that they ought to cling to procedural rules or practices tenaciously even when it prejudices a class of litigants because technicalities must be fetishised at all cost. In the absence of this, the flexibility approach is a solution in search of a problem.

Let us trawl through our legal history to find this spectre of technical reverence. It must be a serious problem indeed for the Court to craft this entire approach in order to deal with this great injustice. Since the Constitutional Court overturned the SCA's judgment, the SCA is as good a place as any to start. Consider the *Telematrix*¹²⁷ decision on the delictual liability of adjudicating bodies for incorrect decisions. Penned by no less a 'formalist' than Harms JA, the Court, when deciding on the exception concerning failure to disclose a cause of action in particulars of claim, said the following:

Exceptions should be dealt with sensibly. They provide a useful mechanism to weed out cases without legal merit. An over-technical approach destroys their utility.¹²⁸

It seems we must go further back in time to find our spectre. Interestingly, Harms JA was not applying precedent from the Constitutional Court inspired by their conception of the 'flexible' approach. He turned instead to the 1960s.¹²⁹

Our courts seem to be as accommodating with respect to how to grant joinders. Consider the problem of whether a principal and an agent may be joined together in the same suit in order to claim relief in the alternative on the merits. This is ideal chum for the rigid, rule-worshipping judge to deny such a joinder if there is no rule providing for this situation. Bale CJ heard precisely this case in 1902 and said:

[I]t is competent for this Court to hold as a matter of convenience in the course of its practice that alternative claims may be allowed under the circumstances, so that the question in dispute between the parties may be determined in one suit, a course which is not only in accordance with modern English practice and with common sense, but one which commends itself to one's *sense of justice*.¹³⁰

So it seems that the rigid procedural approach is not an original sin inherited from the English law. Could it be of Roman-Dutch pedigree? Let us go a step further and consult a Roman jurist, writing long before there was even a concept called the Dutch Republic, and belonging to a society that tolerated moral outrages like slavery and crucifixion as 'just' punishment for sedition. Modestinus tells us:

The reasonableness of law and the mildness of equity quite disallow measures introduced for our benefit being so harshly interpreted as to arrive at severity contrary to our welfare.¹³¹

¹²⁷ *Telematrix (Pty) Ltd v Advertising Standards Authority SA* [2005] ZASCA 73, [2006] 1 All SA 6 (SCA).

¹²⁸ *Ibid* at para 3.

¹²⁹ *Davenport Corner Tea Room (Pty) Ltd v Joubert* 1962 (2) SA 709 (D).

¹³⁰ *F Lenders & Co and FH Lenders & Co, South African Agency v Percy Bros* (1902) 23 Natal Law Reports 231 at 239 (emphasis added).

¹³¹ *Digest* 1:3:25. The translation is by T Gilby in his translation of Thomas Aquinas in his *Summa Theologiae* XXVII at quest.60: art 3.

So if class actions are intended to allow multiple claims to be decided at once, reducing costs for litigants and passing on those burdens to defendants, we cannot introduce and apply certification criteria to thwart this aim. However, we must determine the prior question as to whether it is in a litigant's welfare to have their case and issue be characterised as a class action. Recall Wallis JA's example of a class action over a delictual claim covering negligence, but not damages. Here it is because a litigant may receive less compensation than she is owed if damages are subject to a class action. In any instance, Modestinus quite correctly, establishes the welfare of intended beneficiaries of a law as paramount and so prohibits adopting an unduly severe interpretation that is ambivalent about such welfare. So it seems not even Roman jurists advocate a 'rigid' approach which is at odds with our Constitutional Court's flexible alternative. One starts to wonder whether our courts have constructed a straw man.

But perhaps a different sort of evil has warranted the ire of our highest appellate court. This may be a concern about apartheid judges who sought to apply procedural rules as a moral 'out' aimed at prejudicing the rights of the dispossessed and oppressed. The worry then would be that this hydra may rear another ugly head aimed at marginalising poor, still mostly black, litigants. In the absence of an exhaustive empirical study of most apartheid jurisprudence, let us hone in on the archetype of the apartheid judge who approximates this approach – Chief Justice LC Steyn.¹³² Steyn CJ arguably oversaw the legal institutionalisation of apartheid policies via the interpretation of several key statutes aimed at entrenching racial segregation.¹³³

The best manner of assessing if this rigid approach dominated Steyn's jurisprudence would be to begin with his most convincing critics. In 1982 Edwin Cameron wrote a blistering critique of Steyn's legacy.¹³⁴ What distinguishes Cameron's critique of Steyn's legacy is its thoughtful nuance. Cameron does not do battle with our spectre of procedural rule-worshipping. Instead, he charges Steyn with, amongst other things, what he terms 'executive-mindedness' and a parsimonious sense of justice. Executive-mindedness according to Cameron is an 'excess of ardour in countenancing government power when its exercise is challenged'¹³⁵ and suggests an 'attenuated commitment to protecting the rights of individuals'.¹³⁶ This is not actual bias, but a 'temperamental disposition'.¹³⁷ In a more recent speech in 2010 Justice Cameron, as he is now, warned that the threat from the executive arm of government is no longer a racist state apparatus, but, amongst other things, those within the state who are 'only malignly self-

¹³² Steyn was of course Chief Justice from 1959 to 1971.

¹³³ The most infamous were the Group Areas Act 41 of 1950, the Reservation of Separate Amenities Act 49 of 1953 and the Terrorism Act 83 of 1967. The specific cases which interpreted these statutes during Steyn CJ's tenure include *Cassem en 'n Ander v Oos-Kaapse Komitee van die Groepsgebiederaad en Andere* 1959 (3) SA 651 (A); *Down v Malan NO en Andere* 1960 (2) SA 734 (A) and *Group Areas Development v Hurley* NO 1961 (1) SA 123 (A).

¹³⁴ E Cameron 'Legal Chauvinism, Executive-mindedness and Justice: LC Steyn's Impact on South African Law' (1982) 99 *South African Law Journal* 38.

¹³⁵ *Ibid* at 52.

¹³⁶ *Ibid*.

¹³⁷ *Ibid*.

interested in personal gain'.¹³⁸ Presumably, a temperamental disposition for self-interested state officials would be a modern-day form of executive-mindedness. For our purposes, the validity of this critique is quite irrelevant. We can assume that executive-mindedness, as Cameron describes it, is a cause for deep moral censure for a judge. The deeper question is whether a 'rigid' procedural approach would lead to the evil of executive-mindedness in whatever incarnation.

It does not. Consider the basis for the executive-mindedness theory. Cameron argues that in certain cases concerning the Group Areas Act, Steyn CJ was formalistic and ignored the practical impact of his decisions.¹³⁹ In other cases, Cameron charges Steyn CJ with removing potential procedural and legal impediments to executive 'freedom of action'.¹⁴⁰ Cameron further warns that a judge's insistence on the supremacy of legislative intent allows a judge in covert sympathy with a legislative programme to give full effect to his predispositions without having to accept public responsibility for doing so.¹⁴¹ Equally, Cameron also indicts Steyn CJ for not applying the law of evidence and the procedural rights of detainees in an uncompromising manner even where this would mean a case must be decided against the executive.¹⁴²

At first glance, we may think Cameron contradicts himself. However, a closer reading of his critique reveals that for Cameron it is not that Steyn CJ was a stickler for procedure irrespective of the outcomes of his decisions, but precisely the opposite. Steyn CJ's predisposition was such that he pursued pro-executive results irrespective of the impediments in his way.¹⁴³ So, in some cases procedural rules and guidelines would be a useful cover for his pro-executive leanings. In other cases, procedural rules were an inconvenient, and sadly surmountable, impediment to Steyn CJ's pro-executive sympathies. The problem seems to be a loose, dare we say *flexible*, approach to procedural guarantees, aimed at subordinating individual rights in pursuit of state protection. The lack of consistent and principled application of rules and principles is the mischief. Not the predictability pursued by the rule of law. In fact, a wide discretion unencumbered by mediating guidelines would only serve to expand the space that a pro-executive judge may operate in.

But there is a more fundamental critique to be levelled at Steyn CJ. Cameron argues that Steyn's sense of justice or 'regsgevoel' was lacking in generosity and forbidding in its narrowness.¹⁴⁴ Ponder *Dadoo Ltd and Others v Krugersdorp Municipal Council*.¹⁴⁵ This was a case where the Appellate Division decided whether an

¹³⁸ E Cameron 'Dugard's Moral Critique of Apartheid Judges: Lessons for Today' (2010) 26 *South African Journal on Human Rights* 310, 319.

¹³⁹ Cameron *Legal Chauvinism* (note 134 above) at 55, discussing Steyn CJ's judgment in *Cassem en 'n Ander v Oos-Kaapse Komitee van die Groepsgebiederraad en Andere* 1959 (3) SA 651 (A).

¹⁴⁰ Ibid at 56, discussing Steyn CJ's judgment in *Down v Malan NO en Andere* 1960 (2) SA 734 (A).

¹⁴¹ Ibid at 60.

¹⁴² Ibid at 61.

¹⁴³ Ibid at 59. To quote: 'despite the ambit of judicial choice ... Steyn consistently opted for the executive interpretation.' Cameron also points out at 69 that in legislative interpretation a judge is called upon to affix meaning in accordance with the normative responsibilities of his judicial duties imposed upon him. Presumably Steyn CJ failed to appreciate these normative duties.

¹⁴⁴ Ibid at 62.

¹⁴⁵ 1920 AD 530.

Indian gentleman could acquire property via a company he owned where there was legislation barring Indians from directly owning such land. Innes CJ found that it is a wholesome rule of our law that strict construction be placed upon statutory provisions which interfere with elementary rights.¹⁴⁶ He then found in favour of Dadoo. However, Steyn lamented Innes CJ's construction of the law and decried *Dadoo's* case as evidence of how nothing is more conducive to evasion of the law than clinging to the precise terms of the law.¹⁴⁷ Steyn went further to praise the minority in *Dadoo* as more acceptable to our sense of justice.¹⁴⁸

Here we finally come to the actual moral condemnation of Steyn and his acolytes – their deformed sense of justice. To borrow a line from Toni Morrison:

But don't you understand that the people who do this thing, who practice racism ... are *bereft*. There is something distorted about the psyche. It's like it is a profound neurosis that no-one exams for what it is. It feels crazy. It is crazy.¹⁴⁹

Were we to apply this critique to the s 173 power in the Constitution, we would say that Steyn had a benighted conception of the interests of justice, warped by chauvinism and bigotry. His notions of how to distribute common goods like access to court or of which kinds of people are worthy of corrective relief were bereft.

So, as we examine an apartheid judge's legacy, we can appreciate his brilliance in many areas of law – we do this in philosophy with advocates of slavery like Aristotle and of religious inquisition like Aquinas.¹⁵⁰ But, we do not excuse the application of bereft moral judgement. This is an especially poignant condemnation since the law itself sometimes mandates expressly that such a moral judgement be made by a judge. We must examine this neurosis for what it is.

But in order to level that charge against Steyn or any potential judge with Steyn-like moral deformities we must first determine the ideal or realistic sense of justice of which we speak. An inability to do so subjects litigants to the whims of whatever judge deems a specific category of litigants unworthy of procedural concern.

Hence, the perfectionist notion of distributive justice begins at explicating the obligations of a judge when he is asked to make such a moral judgement. The interests of litigants is paramount. The purpose of these distributive decisions is to subsidise and incentivise good litigation planning for plaintiffs, whilst guarding against abuse of defendants. Mediating guidelines are an indispensable element of these distributive decisions. As is a concern for the nature and ends of the procedural instruments in question.

A conception of justice such as this allows us to question how and when it is in a litigant's welfare to deny or grant a class action. This approach can notionally be applied to other aspects of procedural law such as leave to appeal and

¹⁴⁶ Ibid at 552.

¹⁴⁷ LC Steyn *Die Uittlig van Wette* (5th Edition, 1981) at 46–47. Also discussed by Cameron *Legal Chauvinism* (note 134 above) at 63.

¹⁴⁸ Ibid at 47.

¹⁴⁹ C Rose *An Interview with Toni Morrison* (19 January 1998) Charlie Rose Inc.

¹⁵⁰ See Paul Millett 'Aristotle and Slavery in Athens' 2007 54 *Greece & Rome* 178 and Shadia Drury 'Aquinas and the Inquisition: A Tale of Faith and Politics' (2008) 157 *Salmagundi* 91.

condonation. In any instance, we can no longer be satisfied with vague gestures invoking ‘factors’ in a haphazard manner in how our judicial officers exercise their discretion. It ought to be equally unacceptable that a court would conjure the spectre of the ‘rigid’ judge to shield themselves from scrutiny in the exercise of their discretion. Recall Scalia J’s vivid rebuke of the strategic invocation of a notional spectre:

As to the Court’s invocation of the *Lemon test*: Like some ghoul in a late-night horror movie that repeatedly sits up in its grave and shuffles abroad, after being repeatedly killed and buried, *Lemon* stalks our ... jurisprudence once again The secret of the *Lemon* test’s survival, I think, is that it is so easy to kill. *It is there to scare us (and our audience) when we wish it to do so*, but we can command it to return to the tomb at will. Such a docile and useful monster is worth keeping around, at least in a somnolent state; one never knows when one might need him.¹⁵¹

VI CONCLUSION

The spectre of the ‘rigid’ judge who loves procedural rules for their own sake and is ambivalent about a specific form of justice simply does not and never has existed. Our challenge has always been judges with a stunted understanding of the distributive and corrective demands justice imposes on our courts and use formalism as a stalking horse to conceal a parsimonious conception of justice. We seem to miss that they can only do so as long as we refuse to give content to our moral notions of what ‘justice’ demands. Nonetheless, the ‘rigid’ judge is indeed a useful tool to conjure when a court flouts the rule of law in the exercise of the s 173 discretion, shielding itself from critique by erecting a faux tension between the rule of law and our sense of justice. Our spectre is also useful in dressing down our opponents when they begin to question the legal and moral basis for such an amorphous maybe/maybe not flexible approach.

Fortunately, this cannot obscure the pressing challenge of how our courts can empower prospective litigants who must plan their litigation, whilst facing serious time and resource constraints. This is a serious distributive challenge for our courts. We must be able to hold judges morally accountable in their exercise of judicial power. This then requires we begin to seriously consider what standards of justice we should hold them to. Simply because moral reasoning is difficult and complicated does not mean it can be shirked – especially where the law itself demands it. ‘It depends on the facts and circumstances’ will no longer do. It never did.

¹⁵¹ *Lamb’s Chapel v Center Moriches Union Free School District* 508 US 384 (1993) at 398–399.

Case Comment

No Place for the Poor: The Governance of Removal in *Zulu* and *SAITF*

Irene de Vos*
Dennis Webster†

The law is a terrain of struggle that we cannot avoid.
Zikode, 2012.¹

I INTRODUCTION

In early September 2013, police and Land Invasion Control Unit (Control Unit) officers stormed the Cato Crest informal settlement near Durban and illegally destroyed the homes of many people living there, rendering thousands of them homeless. Later that month, 17-year-old Nqobile Nzuza was shot in the back of the head during protests against the municipality by Abahlali baseMjondolo, the Durban-based shack dwellers movement. Her death brought to three the number of activists who were killed that year in the battle for adequate housing that raged in the settlement.²

Towards the end of the same month, in what proved to be an extraordinarily violent spring for poor people struggling for access and dignity in South African cities, a programme of illegal and forced evictions began in Johannesburg's inner city. Police confiscated goods from informal traders, forcibly removed people from their businesses and dismantled their stalls. The operation was characterised by an excessive use of violence; traders were beaten and assaulted, often whilst running away from the police. Traders were left without their goods and deprived of their businesses and ability to put bread on the table for their families. There were an estimated 30 000 dependants of the evicted traders.³

These two struggles by poor people – for access to land and housing and the use of public space to make a living in South African cities – were taken to

* Senior Legal Researcher and General Counsel, Socio-Economic Rights Institute of South Africa.

† Researcher, Socio-Economic Rights Institute of South Africa.

¹ S Zikode quoted in H Alviar García, K Klare & LA Williams (eds) *Social and Economic Rights in Theory and Practice: Critical Inquiries* (2014) 33.

² Abahlali baseMjondolo 'Marikana Continues: Statement on the Murder of Nqobile Nzuza' (3 October 2013) available at <http://abahlali.org/node/12292/>; R Pithouse 'Durban Poison' (16 October 2013) available at <http://sacsis.org.za/site/article/1817>.

³ G Nicolson & T Lekgowa 'Operation Clean Sweep: Not Just a Clean-up but a Purge of the Poor' (15 November 2013) *Daily Maverick*, available at <http://www.dailymaverick.co.za/article/2013-11-15-operation-clean-sweep-not-just-a-clean-up-but-a-purge-of-the-poor/>; Socio-Economic Rights Institute (SERI) 'The End of the Street?' *Informal Traders' Experiences of Rights and Regulations in Inner City Johannesburg* (2015), available at http://www.seri-sa.org/images/Seri_informal_traders_report_FINAL_FOR_SIGN_OFF_2.pdf.

the Constitutional Court. In *Zulu*, the occupiers of Cato Crest were successful in obtaining an order that disallowed the further destruction of their homes.⁴ Similarly, in *SAITF* the Johannesburg informal traders obtained an order that their eviction from their places of business was unlawful.⁵ The Court victories in both cases came at the end of difficult struggles and failed attempts to engage with local government. Both represent successful strategies of poor people to retain urban land and resources that keep them close to work, social amenities and the benefits of living in a city.

With the Court's roll dominated by socio-economically empowered groups, and in a national context of limited court access for poor people, slow legal processes, and limited *pro bono* services, the rulings also served as a reminder that litigation remains an important avenue for addressing South Africa's extreme levels of poverty and inequality. The state was initially non-compliant with the rulings, however, and their translation into people's lives have been contested, signalling both the obstinance of exclusionary urban agendas and some of the demographic limitations of the constitutional project. In light of criticisms of the Constitution as a document that has entrenched apartheid and colonial legacies,⁶ and of the Court as serving elite interests,⁷ a careful consideration of the two cases is necessary.

In this paper, we do not deal extensively with either of the Court's judgments. Rather, our aim is to interrogate the circumstances and conflicts which led to them being heard in the Court; what Madlingozi has called 'the road from the street to the Court'.⁸ We begin by discussing the two cases in the context of a broader urban crisis in South Africa. We examine issues of urbanisation, housing and the labour market, and suggest that particular forms of governance shape this crisis. In relation to the occupiers of Cato Crest and the Johannesburg informal traders, local government effected forms of governance from which social, spatial and economic justice were by and large excluded. Through a discussion of some of the dimensions of informal work and housing in South Africa, we argue that the municipalities involved, eThekweni (the Municipality) and the City of Johannesburg (the City), conditioned 'states of exception', through which they were able to implement governance distinguished by a pursuit of removal. States of exception are situations which have been structured in such a way that meeting daily needs is made impossible without transgressing the law. The states of exception to which Cato Crest occupiers and Johannesburg informal traders

⁴ *Zulu and Others v eThekweni Municipality and Others* [2014] ZACC 17, 2014 (4) SA 590 (CC), 2014 (8) BCLR 971 (CC).

⁵ *South African Informal Traders Forum and Others v City of Johannesburg and Others; South African National Traders Retail Association v City of Johannesburg and Others* [2014] ZACC 8, 2014 (4) SA 371 (CC), 2014 (6) BCLR 726 (CC).

⁶ JM Modiri 'Law's Poverty' (2015) 18 *Potchefstroom Electronic Law Journal* (argues against notions that the new constitutional dispensation has heralded a 'fundamental rupture and radical break' with colonial and apartheid patterns of racial subordination and socio-economic inequality).

⁷ This critique is more typical among activists and social movements than it is in more elite academic, civil society, and media forums. See for instance T Ngwane as quoted in T Madlingozi 'Social Movements and the Constitutional Court of South Africa' in OV Vieira, U Baxi & F Viljoen (eds) *Transformative Constitutionalism: Comparing the Apex Courts of Brazil, India and South Africa* (2013) 537.

⁸ *Ibid* at 539.

were consigned were contrived and exacerbated by local government, which ‘governed through removal.’ We conclude that, in the face of this urban crisis and an approach to governance characterised by states of exception and removals, the court system remains an important site of struggle for formal recognition of people’s rights.

II SOUTH AFRICA’S URBAN CRISIS

South African cities reflect the broader political and socio-political exclusion experienced by the majority of South African society, which ‘neither benefits from the neoliberal economic model nor is able to influence the political system’.⁹ The country is currently facing an urban crisis characterised by high rates of urbanisation and a lack of affordable accommodation. This is unfolding in the context of a labour market which does not create sufficient formal employment opportunities for people.

The delivery of housing opportunities in post-apartheid South Africa has happened at an impressive rate. Government has created around 3.7 million housing opportunities since 1994,¹⁰ ranging from subsidised free standing houses to the more recent social and rental housing.¹¹ For the better part of this period, housing policy was focused on building large numbers of small Reconstruction and Development Programme (RDP) units to eliminate the apartheid housing backlog and ensure better living conditions for poor people.

Despite these significant achievements, however, there is general consensus among government¹² and civil society that delivery has not been good enough. The housing backlog has risen from an estimated 1.5 million to 2.3 million units.¹³ Other problems have also emerged regarding the RDP programme: a falling rate of delivery, inflated costs, the poor quality of construction, and suspicions of patronage, fraud and corruption in allocation processes.¹⁴ RDP houses are also generally located on urban peripheries where cheap land is readily available. This has effectively re-inscribed apartheid spatial geographies, and has kept poor people far removed from economically viable city centres.

Urbanisation, together with the globalisation and liberalisation of the economy, has been instrumental in the proliferation of urban micro-enterprises, to which some estimates have apportioned 70 per cent of GDP and 80 per cent of urban jobs in sub-Saharan Africa.¹⁵ South Africa’s rapid urbanisation has unfolded in

⁹ J Dugard, T Madlingozi & K Tissington ‘Rights-Compromised or Rights-Savvy? The Use of Rights-Based Strategies to Advance Socio-Economic Struggles by Abahlali baseMjondolo, the South African Shack-Dwellers’ Movement’ in H Alviar García, K Klare & L Williams (eds) *Social and Economic Rights in Theory and Practice: Critical Inquiries* (2014) 23, 27.

¹⁰ The Presidency *Twenty-Year Review: 1994–2014* (2014).

¹¹ SACN (South African Cities Network) *From Housing to Human Settlements: Evolving Perspectives* (2014).

¹² See, for example, the Department of Human Settlements’ *Coherent and Inclusive Approach to Land Policy Framework for Human Settlements* (2016).

¹³ I Turok ‘What will Housing Megaprojects do to our Cities?’ (10 November 2015) available at <http://www.econ3x3.org/article/what-will-housing-megaprojects-do-our-cities>.

¹⁴ SACN (note 11 above).

¹⁵ C Skinner ‘Street-vending in Africa: A Review’ (2008) Working Paper 51, University of KwaZulu-Natal, School of Development Studies.

a context of relatively stagnant growth in formal employment and the increasing casualisation of labour. Macro-economic policies have largely prioritised economic growth over redistribution,¹⁶ overseeing a shift from welfare municipalism to a neoliberal emphasis on fiscal restraint and balancing budgets. Unemployment and inequality have also escalated, and urban poverty is worsening at alarming rates. Recent measures suggest that, when considerations of 'well-being' are taken into account, 62.76 per cent of South Africans are poor.¹⁷ Race remains a key determining factor of these levels of poverty: 70.75 per cent of Black people in South Africa are poor, compared with 56.78 per cent of Coloured people, 20.47 per cent of Asian/Indian people, and only 4.06 per cent of White people.¹⁸

The informal economy has consequently been required to absorb an increasing supply of labour. While this absorption has not been as pronounced as in other African economies, or enough to stem the rampant unemployment in the country, street vending has come to be a dominant activity in South Africa's informal economy, accounting for around 15 per cent of non-agricultural informal employment and almost half of informally self-employed women.¹⁹ While 30 per cent of the estimated 500 000 people making a living on South Africa's streets in 2007 were reported to live in cities, this figure does not account for traders commuting into cities every day to sell their goods and services.²⁰

Excluded from formal job and housing markets, where both the state and market have failed in providing equal levels of access, thousands of people living in South African cities have turned to their own resourcefulness, unofficial networks and arrangements, and informal means to secure some of their most basic needs. For millions of people, shacks are a more adequate housing option than poorly constructed and located government units, which are subject to heavily contested and often corrupt allocation processes.²¹

As Abahlali baseMjondolo have argued, however, this is as much a struggle for basic needs as it is for 'being human' – the demand for houses, land and economic opportunity are expressions of the conditions in which people live. Shacks, and the proliferation of street trade, are popular responses to this deeply entrenched urban crisis. The struggles that led to the judgments in *Zulu* and *SAITF* both arose from within, and as a result of, this urban crisis.

¹⁶ Dugard, Madlingozi & Tissington (note 9 above) at 24.

¹⁷ J Budlender, M Leibbrandt & I Woolard 'South African Poverty Lines: A Review and Two New Money-metric Thresholds' (2015) Working Paper 151, Southern Africa Labour and Development Research Unit 1, 30.

¹⁸ A Finn 'A National Minimum Wage in the Context of the South African Labour Market' (2015) Working Paper 153, Southern Africa Labour and Development Research Unit 1, 5.

¹⁹ G Wills 'South Africa's Informal Economy: A Statistical Profile' (2009) Working Paper 6, WIEGO (Urban Policies).

²⁰ Ibid at 48.

²¹ K Tissington, N Munshi, G Mirugi-Mukundi & E Durojaye *Jumping the Queue, Waiting Lists and other Myths: Perceptions and Practice around Housing Demand and Allocation in South Africa* (2013).

III ZULU AND SAITF

A Zulu

The eThekweni Municipality promised RDP homes to backroom tenants in Lamontville Township in Durban.²² Nothing came of this promise and the tenants, unable to meet their monthly rental obligations, were forced to move from their homes. Unable to afford rent anywhere else, and with the Municipality reneging on its promise to provide homes, the occupiers had nowhere to go. The only place they could move was onto a vacant piece of state-owned land.

As soon as the occupiers arrived, the Municipality demolished their homes.²³ The Municipality was supported by the police and the Municipality's Control Unit. The occupiers' homes were destroyed without a court order. The Municipality argued that it didn't need a court order to destroy homes and evict people, because the occupiers were 'land invaders'. The Control Unit and police regularly visited the property and destroyed their homes. On each occasion, the occupiers rebuilt their homes. Between September 2012 and August 2015, they were evicted and had their homes destroyed 24 times. On each occasion, the demolitions occurred without a court order.²⁴

However, after several evictions, even the police did not agree with the Municipality's conduct of evictions without court orders and informed the Municipality that it would only evict people if the Municipality obtained an eviction order. The Municipality and the MEC for Human Settlements then approached the Durban High Court, but instead of seeking an eviction order, they sought extraordinary relief: an interdict that would apply universally and prohibit any person from occupying several properties, including Lamontville Township. Circumventing the procedural steps provided in the Prevention of Illegal Eviction from and Unlawful Occupation of Land Act (PIE Act),²⁵ the interdict would allow the police and the Municipality to restrain any person from occupying, amongst other properties, the Lamontville Township. The effect of the order was that all the people living on these properties were now doing so unlawfully. Astonishingly, the occupiers themselves were not cited in this application. They were excluded from the court process entirely and this extraordinary relief preventing them from occupying their homes was sought and obtained in their absence. In this way the Municipality effectively obtained an eviction order without following the provisions of the PIE Act and without the occupiers being joined to the proceedings.²⁶

²² *Zulu* (note 4 above) at para 4.

²³ *Ibid* at para 5.

²⁴ *Ibid* at para 6.

²⁵ Act 19 of 1998.

²⁶ The interim order authorised the police and the Municipality to take all reasonable steps to 'prevent any persons from invading and/or occupying' any structures or placing any material on the property, to remove any materials placed on the property and to demolish and/or dismantle any structure that may be constructed on the property after the granting of the order. Specifically Koen J ordered that SAPS and the Municipality may take all reasonable steps:

'interdicting and restraining any persons from invading and/or occupying and/or undertaking the construction of any structures and/or placing of any material upon any of the aforementioned properties.'

A month after the universal interdict was granted, hundreds of the occupiers approached the Durban High Court for relief. They did so in two steps. The first was to seek an order²⁷ restraining the state from demolishing their homes or evicting them without an order of court.²⁸ The second was to request the Court to allow them to be part of the proceedings in which the Municipality and the MEC had obtained the universal interdict. It was this application for leave to intervene which the Durban High Court denied. The Durban High Court²⁹ held that the occupiers did not have a 'direct and substantial' interest in the proceedings. The occupiers took this decision on appeal to the Constitutional Court.

In the Constitutional Court, the occupiers argued that the universal interdict was, in essence, an eviction order. The Municipality argued that it was not an eviction order, nor was it used as an eviction order. Again, the Municipality and the MEC argued that the occupiers did not have a direct and substantial interest in the proceedings.³⁰ The Court held that the universal interdict prevented occupiers from continuing to occupy the property, which amounted to their eviction as it precluded them from either returning to their homes after a temporary absence or because they could be prevented from continuing to occupy the property. This meant that, to this extent, the interim order was an eviction order.³¹ The Court joined the occupiers to the proceedings.³² The majority did not set aside the universal interdict granted by the Durban High Court and referred the matter to the High Court to reconsider. In a minority judgment Van der Westhuizen J set aside the universal interdict. He found that the order could never be legally valid as it had been granted in clear contravention of the PIE Act. The High Court ultimately upheld Van der Westhuizen J's position and dismissed the interdict.³³

B SAITF

As the main arrival point from the rest of the continent, and as the economic heart of a South Africa in which vast sections of the population have little prospect of employment, Johannesburg has developed a thriving informal sector. Nowhere is

²⁷ *J Zulu and 389 Others v Etkekwini Municipality and Others* (unreported) case no 4431/2013.

²⁸ *Zulu* (note 4 above) at para 12.

²⁹ *Ibid* at para 15.

³⁰ They contended that the appellants had no locus standi in the proceedings as the interim order did not affect them or their rights since it only related to invasions or attempted invasions that occurred or would occur after the grant of that order. They submitted that the appellants had already been living on the Lamontville property when the interim order was granted and that the order did not apply to persons who were already in occupation of the Lamontville property when it was granted. In its affidavits and submissions to the Constitutional Court the Municipality was clear: the interim order was not sought for, nor did it apply to, people who were already in occupation of the property. The MEC similarly stated that the occupiers were wrong in their approach to the interim order as it did not cover those who were already in occupation. The legal point was crisp – the interim order does not authorise the eviction of the occupiers as they were in occupation; they therefore had no direct and substantial interest in the case and had no standing. This proved to be false as the state repeatedly used the order to evict occupiers.

³¹ *Zulu* (note 4 above) at para 25.

³² *Ibid*.

³³ *MEC for Human Settlements & Public Works of the Province of KwaZulu-Natal v Etkekwini Municipality and Others; In re: Abahlali BaseMjondolo and Others v Etkekwini Municipality and Another* [2015] ZAKZDHC 69, [2015] 4 All SA 190 (KZD).

this sector more pronounced than on the streets of the inner city, where informal street traders in markets and stalls sell almost everything from muti to make-up, from Nollywood films to fresh fruit and vegetables, from lengths of Mozambican fabric to Nike sneakers, and where most of the inner city eats daily, at Pakistani tea houses, South African chisa nyama stalls and Ethiopian restaurants.

Approximately 7 000 of these traders – the vast majority – were evicted on a mass scale between 30 September and 31 October 2013 as part of the City's 'Operation Clean Sweep'. Accompanied by confiscations of stock, the operation left the traders, many of whom were the breadwinners in their families, without their goods and places of business. The police did not provide a clear or consistent explanation for the evictions. By the end of October, a vast majority of Johannesburg's inner city traders were left without both their places of business and their stock.³⁴ Whilst the traders had been trading 'informally', they were not trading illegally, and were in possession of the necessary permits. Without warning – and without the necessary legal basis – the police arrived and started destroying the traders' stalls. Footage of the action shows massive police brutality. Deprived of a means to support themselves and their families, the traders decided to negotiate with the City, despite legal advice to pursue litigation. The result of these negotiations was an agreement to a 'verification' process. This meant that all the traders would submit to a process in which they would be verified and they would be 're-registered' as informal traders.³⁵ In turn, the City would then allow traders who were verified to return to their stalls. However, the City did not uphold its part of the agreement and the traders were not permitted to return to their stalls after being verified and re-registered. Those who did so were again forcibly removed by the Johannesburg Municipal Police Department (JMPD), who also dismantled the stalls previously used by the traders.³⁶ It became clear that the operation was not an attempt to verify and re-register the lawful informal traders, but was rather a unilateral decision to remove them from the inner city.

Only at this stage did the traders launch an application seeking urgent relief as they had been left without any income for several months. The City argued that their urgency had lapsed as it had negotiated with the traders. The traders ought to have known – argued the City – that the negotiations would not be successful. The Johannesburg High Court agreed with the City and dismissed the application for a lack of urgency. The traders sought direct access to the Constitutional Court. They argued that their eviction was unlawful. The City conceded there was no lawful ground for the operation; however, the City referred to the prejudice that city residents might suffer, 'who no longer have access to ATMs, walk-in banks, cinemas, departmental stores, restaurants and other amenities because of

³⁴ SERI (note 3 above).

³⁵ *SAITF* (note 5 above) at para 8.

³⁶ *Ibid* at para 9.

criminality that hides among the illegal hawkers'.³⁷ The Court dismissed this and held that the eviction had been unlawful.³⁸

IV GOVERNANCE

The current state of local governance, which has developed alongside South Africa's urban crisis, has been characterised as a 'governance deficit', distinguished by corrosive patronage politics, a lack of state engagement with citizens and weak public accountability.³⁹ *Zulu* and *SAITF* are revealing of this deficit, and some other worrying features of urban governance in South Africa. In both cases, local government associated the values of places with people, effectively cutting poor people off from economically competitive city places. Influx control of this kind, characterised by the withdrawal or denial of basic infrastructure and services, is a 'widely deployed means of neoliberal governmentality' and 'has always been an inherent element of modern urban governance'.⁴⁰ Trends of anti-poor approaches have been similarly constant, and amount to governing the poor out of the city.

A Informality: Governance through Exception

State violence against informality and 'unauthorised urbanisation' has a long history in South Africa. Its most notorious manifestation saw informal settlements, which had developed on the peripheries of 'white' apartheid cities presenting the promise of employment, systematically and forcibly removed during the 1960s, 1970s and 1980s.⁴¹ Black South Africans living in cities, and especially those without formal employment, confounded the divisions between tradition and modernity on which colonial and apartheid planning were established and were pathologised as dangerous and 'matter out of place'.⁴²

Ideas that poor black people represent a threat to the urban order persist in the current state's treatment of certain populations. Local government characterised informal homes and work as social and spatial pathologies in both *Zulu* and *SAITF*. In a vocabulary of 'epidemics',⁴³ which has played a crucial role in casting informal settlements as social threats that need to be eradicated, the occupiers in *Zulu* were referred to as 'land invaders'. In *SAITF*, street traders were referenced as the authors of 'crime and grime' in Johannesburg's inner city, which

³⁷ Ibid at para 32.

³⁸ The full order granted was:

- a) Pending the determination of Part B of the application in the High Court, the first to fifth respondents are interdicted from interfering with the trading of the applicants listed in Annexures A and B to this order at the locations they occupied immediately before their removal between 30 September and 31 October 2013.
- b) The first to fifth respondents are directed to pay the applicants' costs in this Court and in the High Court including, in each case, the costs of two counsel.

³⁹ M Van Donk 'Citizenship as Becoming' in *State of Local Governance – Active Citizenship Matters* (2012) 12, 15–17.

⁴⁰ A Selmeczi "'From shack to the Constitutional Court": The Litigious Disruption of Governing Global Cities' (2011) 7 *Utrecht Law Review* 60, 64–66.

⁴¹ J Ferguson 'Formalities of Poverty: Thinking about Social Assistance in Neoliberal South Africa' (2007) 50 *African Studies* 71.

⁴² M Douglas *Purity and Danger: An Analysis of the Concepts of Pollution and Taboo* (1966).

⁴³ N Gibson *Fanonian Practices in South Africa: From Steve Biko to Abahlali baseMjondolo* (2011) 152.

ostensibly results from a proliferation of unregistered trading, despite the City being complicit in this proliferation⁴⁴ and failing in its responsibility to maintain the public environment in the inner city.⁴⁵ Urban planning apparatuses were accordingly deployed ‘to correct or eliminate’⁴⁶ them.

This must be seen in the context of a legal and policy framework which not only recognises those who live and work informally, but also creates rights to protect them. In relation to housing, and within the current context of a limited provision of formal housing, and the constraints of the economy, there is an emerging appreciation that informal work and informal housing are becoming increasingly important to the composition and organisation of cities. This is reflected in the amount of high-level policies which are explicitly designed to manage informality.⁴⁷ Similarly, in the context of informal trade, while a national policy dealing explicitly with informal work remains elusive, local policy acknowledges the crucial economic contributions of informal trade and prescribes a framework of recognition, inclusion and regulation.⁴⁸

Despite national policy and legally binding programmes specifically focused on the recognition and regulation of informal settlements, the approach of local governments to informal settlements has largely been one of relocation, eviction

⁴⁴ C Benit-Gbaffou *In Quest for Sustainable Models of Street Trading Management: Lessons for Johannesburg post Operation Clean Sweep* (2015).

⁴⁵ SERI (note 3 above).

⁴⁶ AY Kamete ‘Missing the Point? Urban Planning and the Normalisation of “Pathological” Spaces in Southern Africa’ (2013) 38 *Transactions of the Institute of British Geographers* 639.

⁴⁷ *Breaking New Ground: A Comprehensive Plan for the Development of Sustainable Human Settlements* (2004) available at http://abahlali.org/files/Breaking%20new%20ground%20New_Housing_Plan_Cabinet_approved_version.pdf. (This policy has held the promise since 2004 of support based state housing interventions, which begin with securing tenure and then enabling governments to upgrade informal settlements in situ in partnership with communities. Informal settlement upgrading before eviction or relocation is also prioritised in the Upgrading of Informal Settlements Programme (UISP).)

⁴⁸ In Johannesburg, the Growth and Development Strategy 2040 (GDS) and the Inner City Transformation Roadmap, which is designed to oversee GDS’s implementation in the inner city, are both decidedly progressive policy frameworks. The GDS, for instance, contains an inclusive conception of the informal sector:

A robust informal sector is essential in supporting economic resilience.... Regulation and policy that manages informality, without destroying informal economic activities and the opportunities they present, serves as additional support, growing resilience further within these economies. The Inner City Transformation Roadmap holds the promise, in policy at least, for a more inclusive approach to the management of informal trade in Johannesburg. It outlines a progressive strategy for the inner city:

A diverse offering of economic activity will be promoted within a mixed economy... Informal trading will be supported within managed linear and demarcated markets and will be integrated with transportation routes and movement plans for the inner city. Micro trading and informal trading will be appropriately regulated... Employment, through meaningful work and livelihood opportunities, will be increased.

The policy also highlights that the issues of ‘crime and grime’ in the public environment of the inner city are best addressed through an urban management approach. With regard to informal trade, it highlights the need for effective management to bolster livelihood opportunities: ‘These issues [“crime and grime”] require effective urban management, ongoing maintenance, and repair by many City departments and entities’, and ‘Well managed informal trade is important both for the dynamism of that sector and for its co-existence with other forms of retail and with pedestrian activity in the inner city.’ The Roadmap reads as a basis for cooperative interventions and management, enabling traders to generate a livelihood in the inner city.

and demolition.⁴⁹ This was expressed most brazenly in the ‘eradication of slums’ discourse that characterised Lindiwe Sisulu’s first term as Minister of Housing.⁵⁰

In *Zulu*, the Municipality’s governance was characterised by the creation of a state of exception. The Municipality promised to deliver RDP homes to occupiers at Cato Crest. It later excluded the occupiers from these benefits on the basis that they were ‘backroom dwellers’ – an exclusion which has no basis in the legal framework. Nonetheless, the Municipality used the backroom-dweller category to create a state of exception from the protection of the legal framework, and thereby exclude occupiers from benefitting from RDP homes. By categorising the occupiers as backroom dwellers, the Municipality made it impossible for the occupiers to bring themselves within the confines of the law or regularise their presence on the property, or any other property for that matter.

The state of exception extended further, however. Once evicted from their backrooms, the occupiers moved onto state-owned land without any consent: they were the very definition of unlawful occupiers.⁵¹ As unlawful occupiers they were entitled to procedural and substantive protection under the PIE Act. This protection means that they had to be given additional specific notice before the eviction application was launched,⁵² their personal circumstances needed to be put before a court, and they could only be evicted if the court deemed it just and equitable.⁵³ Vitaly, if the eviction would lead to homelessness, the state is under an obligation to provide them with alternative accommodation.⁵⁴ Again, the state avoided the legal provisions and its obligations by creating a further state of exception: the occupiers were excluded from their protection under the PIE Act as they were not ‘unlawful occupiers’ but ‘land invaders’. After categorising the occupiers as ‘land invaders’, the Municipality excluded them from the protection they were legally entitled to and the laws that could regulate or legalise their presence on the property. Wherever the occupiers went next, they would be ‘land invaders’, excluded from legal protection and unable to bring themselves within the prescript of the law.

The spatial effect of this sort of governance is that occupiers excluded from legal protection are further excluded from our towns and cities. The occupiers are not welcomed and included, despite legal prescripts that mandate the state to do so, but are relegated to unseen spaces on the urban periphery. Instead of enjoying the benefits of inclusion and protection, they are excluded and wished away.

Operation Clean Sweep flew similarly in the face of the urban-management approach outlined in legal and policy frameworks. In *SAITF*, the City contended that if the applicants were allowed to return to their trading stalls, the inner city

⁴⁹ M Clark & K Tissington ‘Courts as a Site of Struggle for Informal Settlement Upgrading in South Africa’ in L Cirolia, T Gorgens, M van Donk, W Smit & S Drimie (eds) *Upgrading Informal Settlements in South Africa: Pursuing a Partnership-Based Approach* (forthcoming).

⁵⁰ The Housing portfolio has since been changed to Human Settlements.

⁵¹ The PIE Act defines an unlawful occupier as ‘a person who occupies land without the express or tacit consent of the owner or person in charge, or without any other right in law to occupy such land’.

⁵² PIE Act s 4(2).

⁵³ PIE Act s 4(7).

⁵⁴ *City of Johannesburg Metropolitan Municipality v Blue Moonlight Properties 39 (Pty) Ltd and Another* [2011] ZACC 33, 2012 (2) SA 104 (CC), 2012 (2) BCLR 150 (CC).

would be ‘chaotic, uncontrolled and illegal trading with its concomitant crime and grime [will] be permitted to return to the streets of Johannesburg’.⁵⁵ Underlying the City’s stance was the premise that the informal traders and their wares were dirty, which necessitated their removal. Despite the City’s duty to provide a clean environment for traders, it unlawfully⁵⁶ evicted them for contributing to ‘grime’ in the inner city. The traders were not seen as rights bearers, but as nuisances dirtying the city.

The Court, however, recognised the growing reliance on informal sources of income in contemporary South Africa. It acknowledged that the livelihoods of thousands of Johannesburg families depended on informal street trade and that the kind of economic exclusion imposed by Operation Clean Sweep constituted ‘financially perilous conditions’ for informal street traders that traders may not be able to survive.⁵⁷ The Court effectively embraced the struggle for access to city spaces for the purposes of making a living into the ambit of constitutional rights by ruling that ‘the ability of people to earn money and support themselves and their families is an important component of the right to human dignity’, and by adjudging exclusionary urban cleaning operations like Clean Sweep to have ‘a direct and on-going bearing on the rights of children, including their direct rights to basic nutrition, shelter and basic health care services’.⁵⁸

Discourses of urban development, regeneration and improvement have, however, ensured that (often violent) exclusion remains the prevailing experience of poor people living in South Africa.⁵⁹ Solutions to the urban crisis are now typically seen, often through the use of the language of global competitiveness,⁶⁰ as a matter of efficient, bold and creative management that can produce an enabling and secure environment for investment, tourism and entrepreneurship. By defining groups of people outside of the law in both cases, the state’s violence against them, despite being explicitly illegal, was made to appear more legitimate. This is symptomatic of a broader depiction by the South African state of popular practice as criminal, which prepares the grounds on which state violence becomes socially authorised.⁶¹

⁵⁵ *SAITF* (note 5 above) at para 32.

⁵⁶ The unlawfulness of this conduct was never disputed. The Court held that there was no dispute over the entitlement of the traders to trade in the stalls the City allocated to them:

The traders have clear, undisputed rights under section 4 of the By-Laws to do business in the locations where they traded before they were removed. ... The City’s decision to declare certain areas as prohibited or restricted was not made in accordance with the procedure in section 6A(2)(a) of the Businesses Act. The City confesses to this flaw. ... The City has not identified any lawful ground that permits it to frustrate the enjoyment of these rights.

Ibid at paras 25, 27 and 28. The Court expressed concern that the City had described the eviction of several thousand informal traders as ‘convenient’ and instead characterised Operation Clean Sweep as ‘indiscriminate’ and ‘flawed’, finding that the City had ‘gone about achieving its objectives in flagrant disregard of the traders’ rights’.

⁵⁷ *SAITF* (note 5 above) at paras 29–30.

⁵⁸ *Ibid* at para 31.

⁵⁹ S Wilson ‘Planning for Inclusion in South Africa: The State’s Duty to Prevent Homelessness and the Potential of “Meaningful Engagement”’ (2011) 22 *Urban Forum Volume* 265.

⁶⁰ The City of Johannesburg’s aspirational discourse of a ‘world-class African city’ is an example.

⁶¹ R Pithouse ‘An Urban commons? Notes from South Africa’ (2014) 49 *Community Development Journal* 31.

The formality binary has been criticised at length. Keith Hart, who first coined ‘the informal sector’ in 1973, has recently argued that the demise of national capitalism and the rise of neoliberalism since the 1970s has seen the informal economy ‘take over the world’, usually under cover of the rhetoric of ‘free markets’.⁶² But it is important here in order to stress the point that informality is not descriptive of a sector: it is not the shadow of formality. Neither should informality be understood as the object of state regulation, but rather as something brought about by the state itself.⁶³ The state, especially through its legal and planning apparatus, determines what is informal, and in turn, through the deployment of police and private security companies, which kinds of informality will flourish, and which kinds will be torn down and removed. Urban informality, as we see it here, is then a state of exception determined by city planning and development agendas and apparatuses, and is constituted by ‘population[s] whose very livelihood or habitation involve violation of the law’.⁶⁴

B An Unlawful State: The Pursuit of Removal

After excluding the ‘land invaders’ in *Zulu*, and those guilty of ‘crime and grime’ in *SAITF*, from the reach of rights – despite the legal and policy framework designed to include them – occupiers and traders were removed from their homes and places of business respectively.

The intention that underscored the state of exception through which Operation Clean Sweep was achieved was evident from JMPD spokesperson Edna Mamonyane’s announcement at the time that the operation was a success, claiming that the ‘nightmare’ of Johannesburg’s streets had been overcome as ‘now [they] look clean’.⁶⁵ The initiative sought to remove traders permanently from their trading stalls and relocate them to unknown ‘alternative designated areas’, and prohibit them from trading in the interim. These locations would not be close to the high foot traffic in the inner city that the traders depend on for their business. The motivation for the operation was, as the name foreshadowed, one of removal: an attempt to ‘clean’ the inner city.

In *Zulu*, a similar motive of removal informed the Municipality’s actions. And so, when the Control Unit left, the occupiers, having nowhere else to go, rebuilt their homes. The Control Unit regularly visited the property to destroy the homes of occupiers who had rebuilt them after previous destructions. The Municipality had no regard for where the occupiers ought to go to, and sought only to remove them from the land.

The decision to exclude, and the motive to remove, were apparent in the Court proceedings. Up until the Court ruled that the occupiers did have a direct and substantial interest, they had been entirely excluded from the legal process through

⁶² K Hart ‘How the Informal Economy took over the World’ in P Moertenboeck, H Mooshammer, T Cruz & F Forman (eds) *Informal Market Worlds Reader: The Architecture of Economic Pressure* (2015) 33.

⁶³ A Roy ‘Urban Informality: Toward an Epistemology of Planning’ (2005) 71 *Journal of the American Planning Association* 147.

⁶⁴ P Chatterjee *The Politics of the Governed Reflections on Popular Politics in Most of the World* (2004) 40.

⁶⁵ SERI (note 3 above) at 15.

which their removal was sought and their rights to housing and protection against unlawful eviction denied.

The pursuit of removal in *Zulu* reached a point of absurdity. Occupiers, labelled as ‘land invaders’, were evicted without a court order, rendering them homeless. Having nowhere to go they, were forced to ‘invade’ land again, where the eviction cycle began again. In this way, the politics of removal created a chronic eviction cycle in which the occupiers became poorer with each eviction.

V CONSTITUTION AS A SITE OF STRUGGLE FOR POLITICAL SOCIETY

The Constitution imagines all members of South African society as equal, rights-bearing citizens. In the constitutional sense, however, shack dwellers and street traders are only tenuously rights-bearing. While not beyond the reach of the state, which is responsible for them as a population within its jurisdiction, their relationship to it does not always reflect what is imagined in the Constitution.⁶⁶ *Zulu* and *SAITF* express the demographic limitations of the constitutional project. Street traders in Johannesburg and shack dwellers in Durban both occupy a space in what Chatterjee⁶⁷ has termed ‘political society’: they live in informal conditions, whether with regard to access to housing, basic infrastructure and services, or employment, and are compelled to make legal transgressions in their daily lives.⁶⁸

It is important to note that legal proceedings can be ‘long, procedural and expensive affairs that can drain the resources and challenge the integrity’⁶⁹ of any organised and mobilised group of poor people. Litigation, if used at all, only ever constitutes a part of the tactics of these groups. Indeed, the claims to rights are extraordinary in urban struggles in the global south. Everyday survival in ‘political society’ is often more reliant on social networks and local informal arrangements. We focus here only on such extraordinary claims, and indeed only still the litigious hemisphere of these claims. The courts are only a very limited part of disruptive claims to rights, which include mobilisations and protests.

However, that focus is an important one. *SAITF* and *Zulu* constitute moments when distinct groups of South Africa’s ‘political society’ were able to ‘move beyond a politics of invisibility’⁷⁰ and ‘silent encroachment’ and actualise their citizenship in terms of the rights guaranteed them in the post-1994 constitutional framework. Successful struggles in the Court saw them gain ingress to the rights-bearing citizenship from which the state’s governance of removal had excluded them. The Constitutional Court in this context represents a useful site to distil what are often complex and multi-faceted South African urban struggles. In both *SAITF* and *Zulu*, poor people excluded, eventually by force, from prevailing planning agendas made the choice to enlist the courts in their struggle.

⁶⁶ Chatterjee (note 64 above) at 38.

⁶⁷ Ibid.

⁶⁸ C Benit-Gbaffou & S Oldfield ‘Claiming ‘Rights’ in the African City: Popular Mobilisation and the Politics of Informality’ in S Parnel & S Oldfield (eds) *The Routledge Handbook on Cities of the Global South* (2015) 282.

⁶⁹ N Gibson (note 43 above) at 158.

⁷⁰ Benit-Gbaffou & Oldfield (note 68 above) at 283.

Despite the successes achieved in the Court, the struggle of the traders and occupiers continue. For example, in *SAITF* the Constitutional Court's order was not complied with. Not only was the order not complied with, but Nomzamo Zondo (an attorney acting for the traders) was arrested for addressing police. She was arrested only a few hours after the Court interdicted the City from interfering with lawful informal traders' rights to trade.

The JMPD, also a party to the urgent Court application, disregarded the order handed down by Moseneke ACJ and prevented the informal traders from setting up their stalls in the places they occupied prior to Operation Clean Sweep, as authorised by the court order. Zondo was called to the scene and, according to a number of eyewitnesses, was assaulted when she tried to explain to the JMPD officials what the order meant – that the traders were lawfully allowed to trade. She was then arrested and taken to Johannesburg Central Police Station. The exact charges and reasons for her arrest were never made clear as she was never formally charged, only processed. She was detained for over five hours after she was arrested and was denied police bail. It was only on the following day, under threat of contempt and after legal representatives had repeatedly engaged with JMPD officials and showed them the court order, that the JMPD complied with the Constitutional Court's order.

In *Zulu*, the hearing took place on 12 February 2014. The day after the hearing, 13 February 2014, the Municipality demolished a number of structures on the Lamontville property.⁷¹ Despite consistently assuring the Court, in affidavits and submissions, that the interim order would not be used to evict occupiers, the Municipality did so the very next day. This much was admitted by the Municipality in affidavits filed in the Court:

There is an inconsistency between the Municipality's stance on the interim order before this Court prior to and on 12 February 2014 and its reliance upon that order in carrying out the demolitions of 13 February 2014. The Municipality has taken two contradictory positions on the interim order in this matter. Having taken the stance that the Municipality took at the hearing, it was totally unacceptable that the day after the hearing it took a contrary position and carried out the demolitions that it did.⁷²

Aside from the unlawful evictions in the face of undertakings not to evict, the MEC continued to defend an order that operates universally allowing indiscriminate evictions. In addition, the MEC has subsequently sought to 'prevent land invasion' by seeking similar orders. At present there is a universal interdict similar to the one granted by the Durban High Court in *Zulu* operating on a large portion of state-owned land in Durban. It remains to be seen if people will be evicted based on this order as they were in *Zulu*.

VI CONCLUSION

Approaches to governance that centre on the removal of the urban poor, and the interventions they engender, continue to have considerable force in the lives of many people living and working in South African cities. Indeed, both the

⁷¹ *Zulu* (note 4 above) at para 30.

⁷² *Ibid* at para 36.

City of Johannesburg and the eThekweni Municipality are pursuing, through different means, the relocation of Johannesburg informal traders and the eviction of Durban shack dwellers. In Johannesburg, the City has recast its programme as the 'designation of new trading areas', which still amounts to a non-consultative removal of traders,⁷³ while the eThekweni Municipality is currently in the process of refining the interdicts it seeks to use to effect the mass evictions of shack dwellers.

The governance of removal persists beyond these cases as well. In Johannesburg in early 2015, for instance, the City began Operation Ke Molao. The operation, which is still ongoing and echoes much of the logic operationalised during Clean Sweep, is designed to rid the city's traffic intersections of window-washers, peddlers and beggars. Recent statements by the eThekweni Municipality regarding the development plans for Durban in the lead-up to the 2022 Commonwealth Games are similarly worrying. Utilising the now-familiar discourse of 'crime and grime', the municipality is pursuing aesthetic development in its promises to provide a clean and orderly city centre in the next few years.

If these currents persist, the South African urban crisis is unlikely to reach any democratic conclusion, and urban governance will become increasingly exclusionary, authoritarian and repressive.

⁷³ SERI (note 3 above).

