

The introduction of a statutory crime to address third-party foetal violence*

Camilla Pickles

LLB LLM

Academic Associate, Department of Procedural Law
University of Pretoria

OPSOMMING

Die invoering van 'n statutêre misdryf om derdeparty fetale geweld aan te spreek

In *S v Mshumpa* word dit opnuut beklemtoon dat die fetus voor geboorte nie die draer is van grondwetlike regte nie en gevolglik is daar dus ook geen strafregtelike remedies beskikbaar vir die geval waar derdeparty geweld spesifiek gemik is op die beëindiging van swangerskap nie. Die artikel huldig die standpunt dat hierdie situasie aangespreek kan word deur die skepping van 'n statutêre misdryf gebaseer op vroulike reproduksieregte en die reg op menswaardigheid. Sodanige regshervorming sal in lyn met die Grondwet moet wees. Derhalwe kan die fetus steeds nie regstatus verkry nie en mag die misdryf nie in konflik wees met die riglyne vervat in die Wet op Keuse oor die Beëindiging van Swangerskap nie. Die Verenigde State van Amerika is gebruik in 'n vergelykende studie aangesien die kwessie van derdeparty geweld gemik teen die fetus reeds die aandag van die wetgewer en akademiese navorsing in dié jurisdiksie geniet het.

1 INTRODUCTION

*S v Mshumpa*¹ dealt with the controversial issue of violence by a third party aimed at terminating prenatal life. The decision of the Eastern Cape Division emphasised that, until live birth, a foetus is not a legal subject with constitutional rights. As a result of this, a foetus cannot be the victim of criminal conduct. The court refused to develop the common law crime of murder to include a foetus and referred the problem to the legislature to address.

Concerns raised in the article and highlighted by *Mshumpa*, relate to the most effective method of law reform and to the implications of law reform for established legal principles concerning legal subjectivity, the vesting of constitutional rights, and female reproductive rights. In order to avoid these issues, the introduction of a statutory crime is suggested as the preferred method for law reform.

Appreciating the lack of a criminal remedy in relation to a woman's lost pregnancy as a result of violent conduct, this article aims to demonstrate that there is no need to develop or extend definitions of existing crimes, such as murder, to

* The article is based on sections from the author's LLM dissertation entitled *S v Mshumpa: A time for law reform* (UP 2011). It is also based on, and is an adaptation of, an unpublished paper, "The introduction of a statutory crime to tackle third party foetal violence in South Africa", presented at a conference of the Society of Law Teachers of Southern Africa (January 2011).

1 2008 1 SACR 126 (E).

include a foetus. Furthermore, it is not required that a foetus be vested with constitutional rights prior to live birth. The proposed statutory crime is based on constitutionally-sound recommendations and strives to take into account the unique situation of a foetus that is not born alive and consequently is without constitutional rights.²

At first glance, there appears to be a contradiction between a move towards criminalising third-party conduct that terminates prenatal life and the Choice on Termination of Pregnancy Act,³ which provides for the lawful termination of a pregnancy. The article demonstrates that the required law reform is possible without imposing on female reproductive rights. Instead the selected approach uses female reproductive rights as the ground to justify reform.

The legal position in the United States of America is discussed in the article's comparative study. The article examines both federal and selected states' laws in order to obtain a holistic overview of the effectiveness and legitimacy of foetal homicide laws⁴ in the light of female reproductive rights as set out in *Roe v Wade*.⁵

2 FOETAL INTERESTS IN SOUTH AFRICA

2.1 *S v Mshumpa* and the common law crime of murder

Mshumpa is central in demonstrating the need for law reform through the introduction of a statutory crime which will satisfactorily address third-party violence that terminates a pregnancy.

The first accused, Mshumpa, and the second accused, Best, plotted to have Best's pregnant girlfriend, Shelver, shot in the abdomen under the guise of a hi-jacking.⁶ As a consequence of these gunshot wounds, the 38-week-old foetus was stillborn.⁷ The state showed, through the use of expert medical evidence, that prior to the incident, the foetus was viable.⁸ It was proven further that, as a result of being shot, the foetus had tried to breathe in reaction to the pain, causing it to drown in its own blood and amniotic fluid.⁹ Sections of the foetus's cervical spine had also been shattered.¹⁰ Both Mshumpa and Best were charged with murder of the foetus and with the attempted murder of Shelver.¹¹

Murder is defined as unlawfully and intentionally causing the death of another person.¹² In light of this definition, the court was faced with two legal questions:

2 The scope of the article does not extend to those circumstances where a pregnant woman engages in conduct that could terminate her own viable pregnancy. The article is limited to circumstances where a third party causes the termination of a pregnancy through the infliction of violence.

3 Act 92 of 1996 (hereafter "the Choice Act").

4 When discussing third-party foetal violence in the context of the United States, the term "foetal homicide laws" will be used to describe both foetal homicide and feticide.

5 93 SCt 705 (1975).

6 134E.

7 133H.

8 148D. The foetus was viable in that, had it been born on the day of the incident, it would have been able to survive independently from its mother.

9 *Ibid.*

10 133H.

11 134B. Various other charges were filed but are irrelevant here.

12 Burchell and Milton *Principles of criminal law* (2005) 667; Snyman *Criminal law* (2008) 447.

Did the conduct of Mshumpa and Best fall within the ambit of the common law crime of murder? If not, was the court in a position to develop the common law definition of murder to include the termination of prenatal life as a result of third-party violence?

First and foremost, Mshumpa's and Best's intention aside, the court found that their conduct fell outside the ambit of the definition of murder, as the person being killed had to have been alive at the time of the murder.¹³ At the time of the incident the foetus had not been born alive.

The court stated that the development of common law crimes must occur incrementally in accordance with the dictates of the Constitution and should not have retrospective effect since this would offend the principle of legality.¹⁴

In a constitutional era, the principles regarding the High Court's duty to develop the common law and the principle of legality are expressly provided for in the Constitution.¹⁵ In terms of sections 8, 39 and 173, the Constitution places a duty on the superior courts to develop the common law in order to bring it in line with foundational constitutional values.¹⁶ On the other hand, it was noted that section 35(3)(l) of the Constitution, which protects fair trial rights, prohibits an accused from being convicted of a crime that did not exist at the time of the accused person's conduct.¹⁷

The court stated that

"[i]t is one thing to develop the common law in civil matters to eradicate patterns of unequal personal, social and economic domination on the ground that these patterns offend against the foundational values of the Constitution, but it is quite another thing to bring about this development in the face of the legality principle explicitly recognised as a fundamental right itself in section 35(3)(l) of the Constitution".¹⁸

The Constitution does not afford a foetus any fundamental rights, including the right to life. The South African courts have never held a foetus to be the bearer of constitutional rights in its unborn state.¹⁹

The principle of legality barred the court from retrospectively declaring the killing of a foetus as constituting murder and it consequently refused to find Mshumpa and Best guilty of murder.²⁰ When the foetus was killed, the acts or omissions of the two accused did not constitute the offence of murder and therefore they could not be guilty of murder.²¹

It was found that a retrospective declaration would introduce a number of practical difficulties.²² These include formulating a reasonably precise definition of murder to include a foetus, issues linked to foetal viability, whether the crime

13 149E.

14 150I–151A.

15 151D.

16 149H.

17 149I–150A.

18 150A.

19 150B.

20 *Ibid.*

21 151E.

22 150G.

should be restricted to third parties, and how the new definition of murder will coexist with the criminal offence and sanction for illegal abortion under the Choice Act.²³

Further, the court found that a prospective declaration was not possible, since a foetus is not an existing class of persons. Consequently, fundamental rights are not at risk. There are no grounds to justify the prospective development of the common law to extend the definition of murder to include a foetus, since a foetus is not a bearer of constitutional rights.

However, in terms of subsections 12(1), (2)(a) and (b) of the Constitution, the right to bodily and psychological integrity, to reproductive rights, to security, and to control over one's body are protected.²⁴ It was found that these rights are infringed upon when a third party violently terminates a pregnancy.

The court found that a failure to develop the law in order to include the killing of an unborn child as murder will not leave such act unpunished, as the conduct may still be punished as part of the offence committed against the pregnant woman, that is, as murder, attempted murder or any other kind of assault on the pregnant woman.²⁵ The aggravation of the assault in the form of harm to the foetus may suitably be taken into consideration at the sentencing stage.²⁶

It was decided that it is unnecessary to develop the common law crime of murder, since a suitable form of punishment may be found within the ambit of the existing crime of assault against the pregnant mother.²⁷ This approach, the court stated, would avoid the difficulties of formulating a precise, extended definition of murder, and would also avoid making punishment dependent on the stage the pregnancy had reached.²⁸

Pragmatically, the court was faced with the difficulty of providing an adequate definition for the extended crime as well as with the uncertainty that a prospective declaration of a new or extended crime by a high court of first instance would create.²⁹

The judgment concludes with the following significant words:

"I am not saying that there is no merit in making the killing of an unborn child a crime, either as part of the crime of murder or as a separate offence, only that in my view the legislature is, as the major engine for law reform . . . better suited to that radical kind of reform than the courts."³⁰

The court relied on *Masiya v Director of Public Prosecutions*,³¹ referring to *Carmichele v Minister of Safety and Security*³² where the Constitutional Court held: "In exercising their powers to develop the common law, Judges should be

23 *Ibid.*

24 151H.

25 151E.

26 *Ibid.*

27 152A.

28 152B.

29 152C.

30 152D.

31 2007 2 SACR 435 (CC).

32 2002 1 SACR 79 (CC).

mindful of the fact that the major engine for law reform should be the Legislature and not the Judiciary.”³³

The reasons advanced by the court for not developing the common law expressly set the scene for how a foetus is viewed in the legal sphere, and also illustrates the extent of the deficiency of existing criminal law principles in addressing the possibility of a separate offence of third-party foetal violence in terms of which a pregnancy is terminated. Consequently, the law does not directly address the loss experienced by a woman who loses a pregnancy, beyond the attack on her body. Although Shelver had lost something she considered to be her child, the law as it stands does not acknowledge this loss.³⁴ A conviction on a charge of attempted murder carrying an increased sentence is not acceptable, and law reform is required to meet the need to acknowledge that something of intrinsic value had been lost.

Mshumpa unequivocally confirmed the rule that a foetus is not a bearer of constitutional rights, and the court itself specifically stated that, if law reform is required, it is the legislature which should step in. Where the legislature does step in, female reproductive rights, the viability of the foetus, and which persons may be included as perpetrators should all be carefully considered, especially in the light of the Constitution.

2.2 Common law crime of abortion

The common law crime of abortion was committed by “any person who, with the object of defeating the ordinary course of gestation, wilfully applies to a pregnant woman any means by which the untimely expulsion of the foetus is affected”.³⁵

According to *S v Kruger*,³⁶ section 2 of the Abortion and Sterilisation Act³⁷ replaced the common law crime of abortion.³⁸ This Act was then in 1996 repealed by the Choice Act in terms of section 11(2)(a). The prohibition of abortion found in section 2 of the Abortion and Sterilisation Act was not included in the Choice Act, and the Choice Act abolished the statutory offence of abortion.³⁹

33 Para 36. Regarding the extent of the ability of the court to develop the common law, the court quoted *R v Salituro* (1992) 8 CRR (2d) 173: “Judges can and should adapt the common law to reflect the changing social, moral and economic fabric of the country. Judges should not be quick to perpetuate rules whose social foundation has long since disappeared. Nonetheless there are significant constraints on the power of the Judiciary to change the law . . . In a constitutional democracy such as ours it is the Legislature and not the courts which has the major responsibility for law reform . . . The Judiciary should confine itself to those incremental changes which are necessary to keep the common law in step with the dynamic and evolving fabric of our society.”

34 See Kruuse “Fetal ‘rights’? The need for a unified approach to the fetus in the context of feticide” 2009 *THRHR* 126 131 where the author cites the *Daily Dispatch* (12 May 2007) as reporting that, after judgment was handed down, Shelver said: “My child is dead. She was my life. Is nobody guilty of murdering her?”

35 Lansdown *et al* (eds) *South African criminal law and procedure* (1957) 1598.

36 1976 3 SA 290 (O).

37 2 of 1975.

38 S 2 provides that an abortion may not be obtained unless in accordance with the provisions of the Act, and s 3 lists the circumstances under which an abortion may be procured.

39 Kruuse 2009 *THRHR* 126 129.

Section 12(2)(a) of the Interpretation Act⁴⁰ addresses the effect of repealed laws. Botha explains that this provision encompasses those circumstances where a law is repealed by legislation and the repealing legislation is then later repealed itself.⁴¹ The originally repealed law does not regain its force as a result of section 12(2)(a).⁴² Accordingly, the common law crime of abortion is no longer available in order to form the basis of a charge in circumstances such as those in *Mshumpa*.⁴³

2.3 Contravention of the Choice Act

The Choice Act was not enacted to address circumstances where consent for the termination of a pregnancy is lacking or where a pregnancy is terminated as a result of unlawful, violent conduct. Rather, the Choice Act is applicable only to those circumstances where a woman has voluntarily decided to terminate her pregnancy and then actively seeks to give effect to that decision.⁴⁴

The preamble to the Choice Act stipulates that, in acknowledgment of a person's constitutional right to make decisions concerning reproduction, the Act recognises that women have the right of access to appropriate health care services and that the decision to have children is fundamental to women's physical, psychological and social health. Further, the preamble recognises that universal access to reproductive health care services includes terminating a pregnancy. Accordingly, the state has the responsibility to provide reproductive health care for all, as well as safe conditions under which the right of choice can be exercised.

The Choice Act becomes applicable only once a pregnant woman has decided to exercise her right of choice and to terminate the pregnancy. It requires medical personnel who provide termination services to act within the parameters of the Choice Act in order to safely achieve the desired result giving effect to the reproductive rights of women.

2.4 State interest in prenatal life

Regardless of the lack of a direct recognition of foetal interests as set out above, below it will be demonstrated that, through the application of female reproductive rights, the state does have an interest in prenatal life based on the constitutional value placed on human dignity. Further, it will be demonstrated that this interest vests at the moment of foetal viability. Thus the aim to close the gap in the law may be achieved through the introduction of a statutory crime based on the state's interest in prenatal life.

40 Act 33 of 1957. S12(2)(a) provides that, where a law repeals any other law, then, unless the contrary appears, the repeal shall not revive anything not in force or existing at the time at which the repeal takes effect.

41 *Statutory interpretation: An introduction for students* (2005) 42.

42 *Ibid.*

43 See Buthelezi and Reddi "Killing with impunity: The story of an unborn child" 2008 *De Jure* 429 for a contrary view where the authors argue that the common law crime of abortion is still available under South African law and that Mshumpa and Best should have also been charged with this crime.

44 See Buthelezi and Reddi 2008 *De Jure* 429 for a contrary view where the authors argue that Mshumpa and Best should have been charged with contravention of the Choice Act in terms of s 2(2) read together with s 10.

2 4 1 *The Choice Act*

Viewed from the perspective of women, the Choice Act plays a role in advancing and limiting constitutional rights. It advances female autonomy by providing a means for women to exercise their reproductive rights, and, at the same time, limits these rights as the pregnancy progresses and foetal viability sets in. This implies that the Choice Act is performing a balancing act. The decision taken to limit female reproductive rights is also a decision taken to recognise foetal interests based on constitutional values. The values invoked to justify limiting the right to terminate a pregnancy may serve to criminalise third-party violence that terminates a pregnancy.

In *Christian Lawyers Association v Minister of Health*,⁴⁵ the court held that the fundamental right to individual self-determination “lies at the very heart and base of the constitutional right” to terminate a pregnancy.⁴⁶ The court found that sections 10, 12(2)(a) and (b), 14 and 27(1)(a) of the Constitution provide the foundation for the right to terminate a pregnancy.⁴⁷ Like all constitutional rights, this right is not absolute and the state has a legitimate role, in the protection of prenatal life as an important value in our society, to limit a woman’s right to terminate her pregnancy.⁴⁸ However, state regulation cannot amount to a denial of the freedom to exercise the right; the limitation must be reasonable and justifiable in an open and democratic society based on human dignity, equality and freedom as required in terms of section 36 of the Constitution.⁴⁹

A clear manifestation of the state’s interest in prenatal life can be found in the Choice Act. Section 2 of the Act sets out the circumstances and the conditions that allow for the termination of a pregnancy. Section 2(1)(a) states that a pregnancy may be terminated upon the request of a woman during the first 12 weeks of gestation. In terms of section 2(1)(b), from the 13th week up to and including the 20th week of gestation, a pregnancy may only be terminated once a woman has consulted a medical practitioner and that medical practitioner is of the opinion: that continued pregnancy would pose a risk to the mother’s physical or mental health; that there is a substantial risk that the foetus would suffer from severe physical or mental abnormality; that the pregnancy is a result of rape or incest; or that continued pregnancy would severely affect the woman’s social or economic circumstances. In terms of section 2(1)(c), a pregnancy that has reached 20 weeks of gestation may only be terminated if a medical practitioner, after consulting with another medical practitioner, is of the opinion that the continued pregnancy will endanger the woman’s life, will result in severe malformation of the foetus or would pose a risk of injury to the foetus.

The grounds used to justify limiting female reproductive rights provide a foundation that can be used to validate law reform through the introduction of a statutory crime. Once the scope of the limitation is determined, it will be possible to identify the boundaries which the statutory crime of third-party foetal violence may not cross. This is significant because, as it stands, the notion of criminalising conduct that terminates prenatal life is in conflict with the principles of the Choice Act, which accommodate the lawful termination of prenatal life.

45 2005 1 SA 509 (T).

46 518A.

47 518D.

48 527D.

49 527E.

2 4 2 *The value of human dignity*

It is accepted that a foetus is not the bearer of constitutional rights. However, Meyerson rightly points out that this is clearly not the end of the matter.⁵⁰ If it were the end of the issue under consideration, states Meyerson, laws would be passed permitting the termination of prenatal life for any reason right up to the moment of birth.⁵¹ Rather, it is the value of human dignity that is under threat, because

“it is hard to deny that the destruction of foetal life, although it violates no constitutionally protected subject’s right to life, nevertheless undermines human dignity. A foetus is not just a bit of human tissue, comparable to something like an appendix. It is a living organism, whose destruction is not a morally trivial matter but something to be regretted”.⁵²

If it is accepted that the value of dignity serves as the ground for the state to limit female reproductive rights in terms of the Choice Act, this raises the vexed issue of why the Choice Act permits early termination on demand for any reason whatsoever, since it can be claimed that early terminations also offend the value of human dignity. Meyerson asserts that these issues are too controversial for the state to translate into law where the foetus is still very underdeveloped.⁵³ Further, the limitation clause prevents the state from limiting rights for intractable disputed reasons, since a person’s own moral judgements do not justify legal interference by the state where a woman decides to exercise her choice to terminate her pregnancy.⁵⁴ However, beyond early prenatal life, once the foetus becomes more developed, it is capable of feeling pain as it approaches viability and, at this point, its destruction becomes less tied to intractably disputed views and the weight to be afforded to human dignity in competition with female reproductive rights becomes less controversial.⁵⁵

To conclude, the state has an interest in prenatal life once the foetus is viable. Even in these circumstances, the state does not completely limit female reproductive rights, since, in terms of the Choice Act a pregnancy may be terminated, provided that the circumstances calling for a termination fall within the ambit of the Act. Consequently, in order to avoid a conflict between female reproductive rights and the statutory crime, law reform must fall within the parameters of the Choice Act and only be applicable once the foetus is viable.

3 FOETAL HOMICIDE LAWS IN THE UNITED STATES

Very little case law and literature are available in South Africa that directly deal with third-party violence that terminates a pregnancy.⁵⁶ After establishing the existence of a state interest in prenatal life, it is necessary to explore what can be done with this interest to effectively address circumstances such as those in *Mshumpa*. This lack of literature requires the search for a solution to move

50 Meyerson “Abortion: The constitutional issues” 1999 *SALJ* 50 55.

51 *Ibid.*

52 50 59.

53 57.

54 50 57–58.

55 *Ibid.*

56 Only two published articles could be sourced that discuss possible solutions for the circumstances found in *Mshumpa*, namely Kruuse 2009 *THRHR* 126 and Buthelezi and Reddi 2008 *De Jure* 429.

beyond South Africa and consider foreign jurisdictions.⁵⁷ Research indicates that the United States is the only country that has taken steps to address third-party foetal violence; consequently, it forms part of a comparative study.

Although the United States follows a codified legal system, the suitability of this comparative study rests on the fact that the United States is a constitutional democracy which recognises that the right to life does not extend to a foetus, and that women have a constitutional right to terminate their pregnancies.⁵⁸ Even with these fundamental principles in place, the United States has introduced foetal homicide laws at state and federal levels.

Furthermore, the implementation of foetal homicide laws at various levels of governance provides an abundance of information concerning the statutory provisions in place and on the criticisms raised. The criticisms lobbied against foetal homicide laws are highly beneficial for the purpose of this article, because they are based on actual statutory provisions currently being implemented with an *ex post facto*-element, and because the various authors critically review the effectiveness of foetal homicide laws.

In accepting that the right to terminate a pregnancy may conflict directly with the notion of criminalising third-party violence that terminates prenatal life, it is necessary to determine the extent of the right to terminate a pregnancy in the light of the United States Supreme Court's decision in *Roe* and to further investigate how federal and state governments have developed the crime concerning the unlawful termination of prenatal life. More importantly, it needs to be determined whether the introduction by the United States of the crime of third-party foetal violence that terminates prenatal life is legitimate and based on sound legal principles, because, if this has been accomplished, it can serve as a valuable guide for any proposed criminal law reform in South Africa.

In the United States, 38 states have foetal homicide laws.⁵⁹ Although acting on the need to criminalise conduct that terminates prenatal life through negligence or intentional conduct, these states do not deal with this issue with a united vision. The dividing lines between states and the federal government relate to the legal means adopted to address third-party foetal violence and whether the state requires the foetus to be viable at the time the perpetrator inflicts violence on the pregnant woman.⁶⁰

Concerning viability, in 22 states foetal homicide laws are applicable from the early stages of pregnancy.⁶¹ To further deepen the divided approach to foetal homicide laws, under a number of state statutes the termination of prenatal life falls under the crime of homicide, while other states have developed a new crime and designated the non-consensual termination of prenatal life as a special type

57 This is in line with s 39(1) of the Constitution, which provides that, when interpreting the Bill of Rights, foreign law may be considered.

58 See para 4.1.1 below.

59 See <http://www.ncsl.org/IssuesResearch/Health/FetalHomicideLaws/tabid/14386/Default.aspx> (accessed 22 September 2010).

60 According to Tsao "Fetal homicide laws: A shield against domestic violence or sword to pierce abortion rights?" 1998 *Hastings Constitutional LQ (HSTCLQ)* 457 464, where viability is not a requirement, the application of feticide statutes is not linked to the life of the foetus but rather to the harm inflicted on the pregnant woman.

61 See <http://www.ncsl.org/IssuesResearch/Health/FetalHomicideLaws/tabid/14386/Default.aspx> (accessed 22 September 2010).

of murder, namely feticide.⁶² Accordingly, once it has been determined how the state has approached the criminalisation of third-party foetal violence, it will then be necessary to determine whether the particular state requires the foetus to be viable or not.⁶³

Given that each state takes a different approach, it is not feasible to discuss each individually. Instead, a representative state is selected from each of the divergent approaches and the approach taken by the federal government is considered.

At state level, third-party foetal violence is addressed as either homicide or as feticide. Regarding homicide, at common law, and in the absence of a statute at the state level, if prenatal life is terminated before birth as a result of third-party violent conduct a crime has not been committed.⁶⁴ If the foetus is born alive and later dies as a result of injuries sustained while *in utero*, the perpetrator's criminal liability is the same as if he or she had killed any other human being.⁶⁵

In Alabama, a person commits homicide if he or she intentionally, knowingly, recklessly or with criminal negligence causes the death of another human being, including an unborn child at any stage of development.⁶⁶

In Indiana, a person who knowingly and intentionally kills a viable foetus commits murder.⁶⁷ In this case, viability means the ability of a foetus to live outside the mother's womb.⁶⁸ If a pregnancy is terminated prior to viability, the crime of feticide becomes applicable.⁶⁹ Feticide is committed when a person knowingly and intentionally terminates a human pregnancy with the intention other than to produce a live birth or to remove a dead foetus.⁷⁰

In cases where feticide is relevant, different definitions have been allocated. Louisiana defines feticide as the killing of an unborn child, and an unborn child is defined as an individual of the human species from fertilisation and implantation until birth.⁷¹

In complete contrast to the definition of feticide found under Louisiana state law, a person commits the crime of feticide, according to the Iowa Code, when he or she intentionally terminates a human pregnancy, with the knowledge and voluntary consent of the pregnant woman, after the end of the second trimester of the pregnancy where death of the foetus results.⁷² In Iowa, the crime of feticide is not concerned with violence or harm inflicted on the pregnant woman,

62 Gimeno *et al* *Corpus juris secundum: A contemporary statement of American law as derived from cases and legislation* (2006) s 71 451.

63 Wasserstrom "Homicide based on killing of unborn child" *American Law Review (ALR)*, originally published 1998, 671 682.

64 *Idem* 671 680.

65 *Ibid.*

66 S 13A.6.1(a)(2) and (3) of the Code of Alabama.

67 S 35.42.1.1(4) of the Indiana Code.

68 S 16.18.2.367.

69 S 35.42.1.6 of the Indiana Code.

70 *Ibid.* This section does not apply to an abortion performed in accordance with specified statutory provisions.

71 S 14.1.32.5.A and s 14.1.2(11) of the Louisiana Revised Statutes Annotated.

72 S 16.707.7(1) of the Iowa Code Annotated. S 16.702.20 defines viability as that stage of foetal development when the life of the foetus can be continued indefinitely outside the womb by natural or artificial life-support systems.

since her consent to the termination is present and the purpose of this section is to criminalise the wilful termination of a pregnancy that has developed into the third trimester. On the other hand, “non-consensual termination or serious injury to a human pregnancy” is criminalised in section 16.707.8. Serious injury to a human pregnancy, relative to the human pregnancy, means disabling mental illness or bodily injury which creates a substantial risk of death or which causes serious permanent disfigurement, or protracted loss or impairment of the function of any bodily member or organ, and includes, but is not limited to, skull and rib fractures and metaphysical fractures of the long bone.⁷³ The crime of non-consensual termination or serious injury to a human pregnancy is characterised by violence or harm inflicted on the pregnant woman.⁷⁴

The federal government passed the Unborn Victims of Violence Act of 2004,⁷⁵ which stems from the need to protect pregnant women from domestic violence and to deter perpetrators from committing violent acts against pregnant women generally.⁷⁶

The Unborn Victims of Violence Act applies only to a specified list of federal crimes and does not apply to states.⁷⁷ The intention behind the Act is to bring the United States Code in line with other states that have foetal homicide laws. It is not intended that the Act will supersede existing state statutes.⁷⁸

Under federal law, a person who causes the death of a “child who is *in utero*” at the time the conduct takes place will be guilty of a separate offence.⁷⁹ An unborn child is defined as a child *in utero*, and a child *in utero* is defined as a member of the *Homo sapiens* species at any stage of development that is carried in the womb.⁸⁰ The perpetrator will receive the same punishment provided under federal law for that conduct as if the unborn child’s mother had been injured or killed.⁸¹ This offence does not require proof that the perpetrator had knowledge, or should have had knowledge, that the woman was pregnant, or that the defendant intended to cause death or bodily injury to the unborn child.⁸² However, if the perpetrator intentionally kills or attempts to kill the unborn child, the perpetrator will be punished as if he or she had intentionally killed or attempted to kill a human being.⁸³ The Act does not permit the prosecution of any person for terminating a pregnancy where the consent of the mother has been obtained or for medically treating a pregnant woman or the unborn child. In addition, it does not permit the prosecution of any woman with respect to her unborn child.⁸⁴

73 S 16.707.8(11). The stage of development of the foetus is not mentioned under this section.

74 See Rig *Iowa practice series: Criminal law: Homicide* 2009 update 3:52 (Abortion and homicide – Non-consensual termination – Serious injury to a human pregnancy).

75 Hereafter referred to as “the UVVA” in the footnotes.

76 Fitzpatrick “Fetal personhood after the Unborn Victims of Violence Act” 2006 *Rutgers LR* 553 555.

77 Bruchs “Clash of competing interests: Can the Unborn Victims of Violence Act and over thirty years of settled abortion law co-exist peacefully?” 2004 *Syracuse LR (SYRLR)* 133 139.

78 *Ibid.*

79 S 1841(a)(1) of the UVVA.

80 S 1841(d).

81 S 1841(2)(A).

82 S 1841(B)(i)(ii).

83 S 1841(C).

84 S 1841(c)(1)–(3).

4 CRITICISMS LOBBIED AGAINST FOETAL HOMICIDE LAWS IN THE UNITED STATES

At first glance, a number of issues may be identified regarding foetal homicide laws. However, the discussion focuses on the Unborn Victims of Violence Act, since this is a federal statute. First, the Act is extremely far-reaching in that a perpetrator need not be aware that the woman is pregnant or even have the intention to terminate the pregnancy or harm the foetus through the infliction of violence. Secondly, foetal viability not being an element raises concerns about adequately proving causation. Thirdly, pro-choice activists are also concerned that the Act extends personhood to a non-viable foetus and that this contradicts *Roe*'s finding that the state's interest only becomes compelling from the point of viability onwards. Fourthly, the Unborn Victims of Violence Act has left its intended track of dealing with violence against pregnant women. These issues have been identified as fundamental flaws in the federal law and in seminal state law.

4.1 Foetal viability

4.1.1 *Female reproductive rights and the extension of personhood to a foetus*

Roe deals with a woman's constitutional right to terminate her pregnancy, and, for purposes of criminal law, it also explicitly sets out the scope of the state's interest in prenatal life.⁸⁵

In *Roe*, the Supreme Court held that Texas laws prohibiting the termination of pregnancies at any stage of development of the foetus, except to save the life of a woman, are unconstitutional.⁸⁶ The court held that the due process clause of the Fourteenth Amendment protects a person's privacy rights against state interference, which includes a woman's right to terminate her pregnancy.⁸⁷ Working on a trimester framework, the court held that, in the first trimester, the decision to terminate a pregnancy is the woman's decision alone.⁸⁸ However, the right to terminate a pregnancy is not absolute, and, as the pregnancy develops into the second trimester, a state may only limit the decision to terminate if it is in the interests of the health of the mother.⁸⁹ The Supreme Court stated that, at the point of viability and onwards, the state's interest becomes sufficiently compelling to sustain the regulation of factors that govern the termination of pregnancies.⁹⁰ Further, *Roe* confirmed that a foetus is not a bearer of constitutional rights and, accordingly, that a foetus does not have a claim to the right to life.⁹¹

In *Planned Parenthood of Southern Pennsylvania v Casey*, the United States Supreme Court rejected *Roe*'s trimester framework and found that the courts should rather employ the undue burden standard to protect the right to terminate

85 Termination of pregnancy laws at state level will not be discussed, because *Roe*, as a Supreme Court decision, is binding on the entire legal system of the United States. See Morrison *Fundamentals of American law* (1996) 54.

86 708.

87 *Ibid.*

88 732.

89 *Ibid.*

90 727.

91 708.

a pregnancy.⁹² The Supreme Court further found that a state has an interest throughout the pregnancy and that, to promote this interest, it may take measures to ensure that a woman's choice is informed, but that these measures cannot place an undue burden on a woman's right to terminate her pregnancy.

Consequently, the trimester framework is no longer applicable as a result of the decision in *Planned Parenthood*, but the right to terminate a pregnancy as a privacy right and the distinction between a viable and a nonviable foetus as decided in *Roe* remain intact.⁹³

Foetal homicide laws that do not require foetal viability are criticised for infringing a woman's constitutional right to terminate her pregnancy, since states illegitimately begin to act on their interest in prenatal life before the foetus is capable of meaningful life outside the womb.⁹⁴ Pro-choice activists are concerned that the courts may interpret the Unborn Victims of Violence Act as a congressional declaration that a foetus is a person from the moment of conception, which could lead to the demise of the *Roe* decision and the demise of termination of pregnancy rights.⁹⁵

Those states that require a foetus to be viable in terms of foetal homicide laws assert that this requirement is dictated by the *Roe* decision.⁹⁶ These foetal homicide laws fall within the boundaries of *Roe* and are in harmony with the constitutional significance of foetal viability that, once a foetus is viable and capable of meaningful life outside the mother's womb, the state's interest in prenatal life develops into a compelling interest and that state interference with the exercise of female autonomy, and in terms of limiting third parties' conduct, is logically and biologically justified.⁹⁷

Bruchs argues that, despite the abortion exception contained in the Unborn Victims of Violence Act, the Act expressly confers legal personhood on both non-viable and viable foetuses, because section 1841(d) defines an unborn child as a child *in utero* at any stage of development.⁹⁸ The termination of a pregnancy was never considered to be the termination of life, because, in the light of *Roe*, a

92 112 SCt 2791 (1992) 2799. Here, a Pennsylvania statute placed restrictions on the right to terminate a pregnancy to the extent that a physician was required to give a woman information advocating childbirth over terminating a pregnancy in order to enable the pregnant woman to give her informed consent, and to wait at least 24 hours after the information had been given before the pregnancy could be terminated. The statute also required the physician to obtain parental consent regarding the termination of a minor's pregnancy and to ensure that a married woman had notified her husband of her decision to terminate her pregnancy. For all termination procedures performed physicians were further required to keep records.

93 Tao 1998 *HSTCLQ* 457 466.

94 457 470.

95 Rosen "A viable solution: Why it makes sense to permit abortion and punish those who kill fetuses" 2003 *Legal Affairs* 20. See also Holzapfel "The right to live, the right to choose, and The Unborn Victims of Violence Act" 2002 *J on Contemporary Health L and Policy* 431 439 where, according to the Planned Parenthood Association of America, it is stated that the UVVA is solely designed with the intention to erode the foundation of *Roe* and that the Act has shifted the focus away from violence against women to violence against a foetus as if it is a legally-recognised person.

96 Tao 1998 *HSTCLQ* 457 467.

97 *Ibid.*

98 Bruchs 2004 *SYRLR* 133 136.

person does not include a foetus.⁹⁹ However, the wording of the Unborn Victims of Violence Act explicitly states that the perpetrator's conduct must cause "the death of . . . a child *in utero*", which implies that the foetus is a live child.¹⁰⁰

Further, Bruchs states that United States abortion law requires a balancing act between life interests and liberty interests under the Fourteenth Amendment.¹⁰¹ On the one hand, the Supreme Court of America in *Roe* recognised a woman's right to privacy and her right to have an abortion as a protected liberty interest under the Fourteenth Amendment, and, on the other, it also recognised the potential life of a foetus to be legitimately protected by the state.¹⁰² The standard set in *Roe* aids in reconciling the differences between the interests of women and the interest in potential life and prevents a situation where one interest is included at the expense of the other.¹⁰³ The balanced relationship is put at risk by the Unborn Victims of Violence Act recognising a complete life interest in prenatal life, which could force a court to choose one interest over the other.¹⁰⁴ When a court is faced with the situation of having to decide which interest is deserving of greater protection, it would almost certainly take the necessary steps to protect the life interest.¹⁰⁵

4.1.2 Causation

Fundamental to criminal liability is the requirement that it must be proven beyond reasonable doubt that the perpetrator's conduct caused the death of the victim.¹⁰⁶

Given the inherent fragility of prenatal life before viability, it is argued that proving causation will be impossible.¹⁰⁷ A non-viable pregnancy is wrought with uncertainties; many women face spontaneous abortions due to a number of factors, including genetic or developmental defects in the foetus, uterine abnormalities, maternal trauma, illnesses, substance abuse or the presence of toxins in the foetal or maternal environment.¹⁰⁸ If any of these factors are present, it is questionable whether it can be successfully proven that the perpetrator's conduct caused the termination of the woman's pregnancy.

In addition, in those cases where the pregnancy terminates hours after contact with the perpetrator, it will be difficult to prove that the termination of the pregnancy was not a result of natural complications,¹⁰⁹ that is, that the inherent fragility of the non-viable foetus was not a *novus actus interveniens*.

In *People v Davis*,¹¹⁰ the Supreme Court of California found that killing a foetus, whether viable or not, constituted murder, provided that it had developed beyond embryonic stage. The defendant, during the course of a robbery, had

99 143.

100 S1841(a)(1).

101 Bruchs 2004 *SYRLR* 133154.

102 *Ibid.*

103 *Ibid.*

104 *Ibid.*

105 Bruchs 2004 *SYRLR* 133 154.

106 Tsao 1998 *HSTCLQ* 457 474.

107 473.

108 *Ibid.*

109 *Ibid.*

110 7 Cal 4th 797 (1994).

assaulted a 23 to 25-week pregnant woman by shooting her in the chest. The woman survived, but a day later gave birth to a stillborn child.¹¹¹

Mosk J raised this very concern.¹¹² The problem with a nonviable foetus and causation is a medical reality, since, by lowering the required gestational age of the foetus to seven weeks, the majority overlook the fact that the more immature the foetus, the more likely it is to die by spontaneous abortion.¹¹³ Mosk J stated that the incidence of spontaneous abortion is generally believed to be 15 to 20 per cent of all pregnancies, but that substantial numbers of spontaneous abortions are unreported or happen very early and sub-clinically and are estimated to be as high as 50 to 78 per cent of all pregnancies.¹¹⁴

The mere fact that a foetus spontaneously aborts itself after a pregnant woman is attacked by a perpetrator does not mean that the perpetrator's action is a substantial factor in causing the termination of the pregnancy.¹¹⁵ Mosk J stated that a prosecutor could mislead the jury into convicting the perpetrator of foetal murder without establishing causation with certainty.

4 2 Lack of intention in terms of the Unborn Victims of Violence Act

According to the Unborn Victims of Violence Act, the prosecution need not prove that the accused, while committing a federal crime, had knowledge, or should have had knowledge, that the victim of the underlying federal offence was pregnant or that the accused even intended to cause the termination of the pregnancy.¹¹⁶

The Unborn Victims of Violence Act requires the application of the doctrine of transferred intent.¹¹⁷ Under this doctrine, a perpetrator who intends to kill one person, but instead kills another, "is deemed to be the author of whatever kind of homicide would have been committed had he killed the intended victim".¹¹⁸

Accordingly, where a perpetrator commits a crime against a pregnant woman and as a result terminates her pregnancy, the unlawful intent is transferred from the pregnant woman to the foetus.¹¹⁹ Once the doctrine of transferred intent has been established, it is necessary to prove beyond reasonable doubt that the woman was pregnant and that the perpetrator's criminal conduct directly caused the harm.¹²⁰

In *Davis* Mosk J, dissenting from the majority, stated that this rule is controversial enough without the added dimension of making a foetus the victim of

111 800.

112 840.

113 *Ibid.*

114 841.

115 *Ibid.*

116 S 1841(B)(i) and (ii).

117 Jolicoeur-Wonnacott "The Unborn Victims of Violence Act: Friend or foe to the unborn?" 2000 *Thomas M Cooley LR (TM Cooley LR)* 563 567.

118 *Ibid.*

119 *Ibid.*

120 *Ibid.* Sepinwall "Defense of others and defenseless 'others'" 2005 *Yale J of L and Feminism* 327 340-341 states that application of transferred intent presupposes that the intended and unintended victims are of equal moral worth and that it makes sense to transfer intent from one to the other, therefore presuming that the unintended victim is the sort of being contemplated by the criminal code as a victim.

murder.¹²¹ The severity of this principle will be considerable if, under the relevant statute, foetal viability does not play a role because the felony murder rule would make such conduct a capital offence for even an invisible victim.¹²²

4.3 Intended victim

To illuminate the fact that foetal homicide laws have wandered from their intended track of tackling violence against pregnant women, Tsao states that such laws are criminal statutes and that perpetrators are prosecuted for ending prenatal life and not for depriving a woman of her reproductive choice to bear children.¹²³

Fitzpatrick argues that the focus should be on the harm caused to the pregnant woman rather than on the harm caused to the foetus.¹²⁴ When the focus is placed on the foetus, the statutory provision is required to consider the foetus as a person in order for it to be a victim of crime.¹²⁵

An alternative to the Unborn Victims of Violence Act considered by Fitzpatrick is the Motherhood Protection Act of 2004,¹²⁶ which focuses on establishing a criminal offence of causing harm to a pregnant woman.¹²⁷ Section 2(a) states that whoever engages in any violent or assaultive conduct against a pregnant woman resulting in the conviction of the perpetrator for a violation of any section provided for in subsection (c),¹²⁸ and who causes an interruption to the normal course of the pregnancy resulting in prenatal injury or termination of a pregnancy, will face further punishment in addition to that imposed in terms of subsection (c). The perpetrator can be sentenced to payment of a fine or to imprisonment for not more than 20 years, or both.¹²⁹ However, if the conduct causes the termination of a pregnancy, the perpetrator could face a fine or any term of imprisonment, including life imprisonment.¹³⁰

Hickcox-Howard states that the Motherhood Protection Act has failed because of the concern that someone who assaults a pregnant woman and causes the termination of her pregnancy will not be punished harshly enough because the attack on the pregnant woman does not result in the pregnant woman's death.¹³¹ Additionally, the loss of a pregnancy demands a state response which is on a par with the state's response to murder. However, the Motherhood Protection Act only takes one victim into account, which is perceived as stripping unborn victims of their dignity.¹³²

121 838.

122 837.

123 1998 *HSTCLQ* 457 470.

124 Fitzpatrick 2006 *Rutgers LR* 553 578.

125 *Ibid.*

126 Hereafter "the MPA".

127 Fitzpatrick 2006 *Rutgers LR* 553 578.

128 All the listed sections relate to federal crimes, which is similar to the approach taken by the UVVA. See Alongi "The Unborn Victims of Violence Act and its impact on reproductive rights" 2008 *Washington and Lee Journal of Civil Rights and Social Justice (WLLCRSJ)* 285 293.

129 S 2(b)(1).

130 S 2(b)(1).

131 "The case of pro-choice participation in drafting foetal homicide laws" 2008 *Texas J of Women and the Law* 317 326.

132 *Ibid.*

The Unborn Victims of Violence Act and the Motherhood Protection Act include two separate crimes. The Motherhood Protection Act does not recognise the foetus as a victim but rather focuses on the unlawful interference with a pregnant woman's constitutional right to choose to continue with her pregnancy.¹³³ As a result of this approach, such Act has successfully steered away from providing a foetus with any kind of legal personhood to the extent that it does not mention any legal subject other than the pregnant woman.¹³⁴

5 PROPOSED LAW REFORM IN SOUTH AFRICA

The intention of the article was to determine whether it is possible to protect a developing foetus from third-party violence. In the light thereof, all efforts were foetus-oriented. The approach adopted in the United States demonstrates that any law reform based solely on the foetus as an entity separate from the pregnant woman will not be feasible. It is not the purpose of the article to introduce controversy surrounding reproductive rights and the right to life. Rather, the law reform proposal aims to introduce a legally-sound statutory crime that prevents the termination of a pregnancy through the infliction of third-party violence and does not conflict with established legal principles.

5.1 Application of foetal homicide laws in South Africa

Having established that the state has a legitimate interest in viable prenatal life, it is necessary to determine whether the United States position concerning foetal homicide laws can provide direction for South Africa. Using the various criticisms lobbied against foetal homicide laws as a guideline, it will be demonstrated that United States foetal homicide laws do not offer a workable solution, especially when measured against the strength of the South African constitutional regime. Less radical statutory provisions can be employed to bring about the same result.

5.1.1 *Foetal viability and the state's interest*

The fact that the Unborn Victims of Violence Act takes effect from the moment of conception will prove fruitless in South Africa. The state's interest is not strong enough prior to viability, and this is very clear from the provisions of the Choice Act discussed above.

There are no grounds available to justify criminalising conduct that terminates a pregnancy before foetal viability. In cases where grounds can be established to justify criminal sanctions, causation requirements stand in the way.¹³⁵

5.1.2 *Culpability*

In South Africa, *S v Bernardus*¹³⁶ rejected the doctrine of transferred intent. Culpability requires that the accused acts negligently or with intention and that culpability and the unlawful act must be contemporaneous.¹³⁷

133 Alongi 2008 *WLLCRSJ* 285 292 294.

134 294.

135 See para 5.1.3 below.

136 1965 3 SA 287 (A). Prior to this decision, culpable homicide was defined as the unlawful causing of another person's death, without any culpability.

137 Burchell and Milton (2005) 544; Snyman (2008) 153.

Linked with the culpability requirement, the accused has a right not to be deprived of his or her freedom arbitrarily or without just cause.¹³⁸ The primary purpose of section 12(1) of the Constitution is to protect a person's physical liberty and to prevent unwarranted intrusion by the state.¹³⁹

Ramraj argues that O'Regan J in *S v Coetzee*¹⁴⁰ expanded the substantive protection of section 12(1) to include a "constitutional doctrine of criminal fault".¹⁴¹ O'Regan stated that the Constitution recognises that the state may not deprive its citizens of freedom for reasons that are not acceptable, and that, in cases where freedom is limited for acceptable reasons, the state must only do so in a manner that is procedurally fair.¹⁴² Accordingly, the Constitution limits the reasons for which the legislature may deprive a person of his or her freedom and the manner in which this is done.¹⁴³

Culpability does not have to be in the form of direct intent, and the law has recognised that even negligence can give rise to criminal liability.¹⁴⁴ O'Regan held that

"it is only when the legislature has clearly abandoned any requirement of culpability or when it has established a level of culpability manifestly inappropriate to the unlawful conduct or potential sentence in question, that a provision may be subject to successful constitutional challenge".¹⁴⁵

Placing the South African principles relevant to culpability in the context of the United States foetal homicide laws, a perpetrator would need to be aware of the pregnancy and have the intention to terminate the pregnancy through the infliction of violence on the pregnant woman. Consequently, the fact that the Unborn Victims of Violence Act does not require the intention to terminate a pregnancy makes the application of this provision impossible in the context of South African criminal law.

5.1.3 Causation

The Unborn Victims of Violence Act attaches criminal liability to the termination of a non-viable foetus. The application of this provision in South Africa would be inconsistent with the accused's right to be presumed innocent in terms of section 35(3)(h) of the Constitution.

The presumption of innocence is necessary to reduce the possibility of wrongful conviction and finds expression in the reasonable doubt standard which places the burden of proof on the prosecution to prove the accused's guilt beyond reasonable doubt.¹⁴⁶ The rationale for the presumption of innocence is to protect an individual's rights from the potentially-coercive authority of the state and also includes policy considerations that maintain the legitimacy of the criminal justice system.¹⁴⁷ The likelihood of an erroneous conviction threatens the legitimacy of

138 S 12(1)(a) of the Constitution.

139 Currie and De Waal *The Bill of Rights handbook* (2005) 292.

140 1997 1 SACR 379 (CC).

141 2002 SAJHR 255 237.

142 473F.

143 438A.

144 443A.

145 443C.

146 Currie and De Waal (2005) 747.

147 Schwikkard "The presumption of innocence: What is it?" 1998 SACJ 396 407.

the criminal justice system and weakens the normative value of criminal law. Placing the burden of proof on the prosecution will keep the frequency of wrongful convictions within tolerable parameters.¹⁴⁸

In the context of proving beyond reasonable doubt that the accused caused the termination of a pregnancy, foetal viability plays a key role. Sarkin-Hughes discusses foetal viability in connection with brain birth.¹⁴⁹ Accepting that brain death is indicative of the end of life, Sarkin-Hughes questions whether brain birth should be the indication of the presence of life.¹⁵⁰ In order for the brain to operate as a functioning organ, neocortical activity is essential, because neocortical cells must be developed for thought, emotion and consciousness to occur.¹⁵¹ At around 24 weeks of gestation, the dendritic spines¹⁵² appear, which are essential for brain circuitry, and, from roughly 19 weeks' gestation, neural connectors between brain cells also appear.¹⁵³ At roughly 22 weeks' gestation, thalamocortical¹⁵⁴ connections develop, which are essential for neocortical reception of bodily sensation.¹⁵⁵ Accordingly, Sarkin-Hughes states that 22 weeks' gestation is important from a brain birth perspective.

Sarkin-Hughes further states that viability is a fixed moment in foetal development, that it does not shift as medical science develops, and has always remained at 22 weeks' gestation.¹⁵⁶

“While there have been advancements in medical science which have been perceived as pushing the viability line further back, what has occurred is that the incidence of survival of foetuses born after viability has increased over the years. No actual change to the viability threshold has occurred. Today additional aid can be given to these 22-week-old infants but in foetuses younger than this the organs have not developed sufficiently; so whatever aid, however technologically advanced, is given it will not assist the foetus to survive.”¹⁵⁷

In the case of a non-viable foetus, vital organs have not developed sufficiently to allow for proper performance of their usual functions and, unless the organs are functioning, life cannot be sustained, even with the aid of medical intervention.¹⁵⁸ Sarkin-Hughes asserts that the cut-off line for foetal viability is at 20 weeks, which takes into account a two-week margin of error that may occur

148 407–408.

149 Sarkin-Hughes “A perspective on abortion legislation in South Africa’s Bill of Rights era” 1993 *THRHR* 83 88.

150 *Ibid.*

151 *Ibid.*

152 Dendritic spines are short outgrowths of neurons that relay electrical impulses to the brain. See http://www.google.co.za/search?hl=en&rlz=1W1PCTC_en&defl=en&q=define:Dendritic+spines&sa=X&ei=uaKyTMegO9GYOqefiPsF&ved=0CBQQkAE (accessed 10 October 2010).

153 Sarkin-Hughes 1993 *THRHR* 83 88.

154 The thalamocortical system relates to the connecting cortex and the thalamus. See <http://wordnetweb.princeton.edu/perl/webwn?s=thalamocortical> (accessed 10 October 2010). The thalamus forms part of the forebrain and transmits nerve impulses up sensory pathways to the cerebral cortex. See <http://www.tuition.com.hk/psychology/t.htm> (accessed 10 October 2010).

155 Sarkin-Hughes 1993 *THRHR* 83 88.

156 89.

157 *Ibid.*

158 *Ibid.*

when estimating the gestational age of the foetus.¹⁵⁹ Another important factor linked to foetal viability relates to the fact that a foetus can feel pain from viability onwards.¹⁶⁰

Sarkin-Hughes's argument is extremely valuable in illustrating the potentially-important role viability can play in the context of proving causation and in meeting the burden of proof required by the state for purposes of third-party violence that terminates a pregnancy.

6 CONCLUSION AND RECOMMENDATIONS

Inadvertently, female reproductive rights provide the solution to the problem faced in *Mshumpa*. Although reproductive rights are the limiting factor to foetal victimhood or personhood, they are also the key element in justifying state intervention where a pregnancy is terminated as a result of third-party violence. It is not the Unborn Victims of Violence Act with its far-reaching and radical provisions that offers a practical solution, but rather something similar to the provisions of the Motherhood Protection Act, focusing on violence against pregnant women. It needs to be noted that a continued pregnancy is a conscious decision to bear children and amounts to a woman exercising her reproductive rights. Where a pregnancy is terminated as a result of unauthorised conduct, it amounts to an infringement of her constitutional right to self-determination. From this perspective, third-party violence that terminates a pregnancy can receive sufficient legal support to justify law reform.

The starting point is to use the state's interest in reproductive rights to establish an entrenched need to prevent violence against pregnant women. This need is heightened by the state's interest in viable prenatal life. Where a third party terminates a pregnancy without the woman's consent, he or she commits a crime against the pregnant woman. This approach keeps the spotlight on the pregnancy and not on prenatal life as a separate consideration. The combination of female reproductive rights and the state's interest in viable prenatal life establishes a solid ground to justify the criminalisation of third-party violence that terminates a pregnancy.

The recommendations that follow are based on the Motherhood Protection Act, with adjustments to allow the proposed crime to be compatible with the South African criminal law principles of culpability and causation. In addition, the recommendations are guided by the criticisms levelled against foetal homicide laws in the United States.

If the focus is to be shifted from the foetus to the pregnant woman, wording similar to that of the Unborn Victims of Violence Act and other foetal homicide laws will have to be abandoned. Since it is not the intention of the proposed crime to advance the right to life, but to deter the violent termination of a pregnancy, phrases like "mother", "unborn child" and "causing the death of the unborn child" will have to be avoided. The statutory provision must be personhood-neutral and refer to the pregnant woman and her pregnancy only.

Foetal viability will be a requirement in terms of the statutory crime for purposes of causation and to legitimately act as an extension of female reproductive rights in South Africa.

159 *Ibid.*

160 Meyerson 1999 *SALJ* 50.

It would be irrational to strictly define foetal viability in terms of a specific number of gestational weeks, because there is a certain amount of uncertainty when trying to determine the exact stage of foetal development. Foetal viability should rather be defined as the moment in foetal development when it can be found that the foetus is capable of living an independent life separate from the pregnant woman, with or without the aid of medical support.

The proposed crime must be an intentional crime requiring the perpetrator to intentionally cause the termination of the pregnancy while committing an underlying common law crime or statutory offence. All forms of intention should be applicable.

In line with these legal specifications, the proposed statutory provision should read as follows: Any person, other than the pregnant woman herself, who engages in unlawful conduct in terms of the common law or statute, and thereby intentionally causes the termination of a viable pregnancy or interrupts the normal course of a pregnancy, including the termination of the pregnancy other than by live birth, is guilty of an offence under this Act. Punishment for that separate offence will be in addition to that of the underlying offence found in common law or in terms of statute.

The article has revealed that the task of addressing third-party violence that terminates a pregnancy is not impossible. Constitutional female reproductive rights offer the most effective solution for law reform in South Africa. Furthermore, the value of dignity that vests in a foetus plays an important supportive role in justifying legislative intervention. In fact, the collective strength of female reproductive rights and the value of dignity demands that the law acknowledges the loss caused by third-party violence that terminates a pregnancy from the perspective of the woman herself, her relatives and society as a whole. This statutory crime goes beyond “determining an appropriate sentence”, as was done in *Mshumpa*,¹⁶¹ and amounts to more than the mere consideration of aggravating factors.

161 152A.

Constitutional damages and proof of damage: Recent contributions of the Privy Council

Chuks Okpaluba

LLB LLM PhD

Adjunct Professor, Nelson Mandela School of Law,
University of Fort Hare

OPSOMMING

Grondwetlike skade en bewys van skade: Onlangse hydraes deur die Geheime Raad

In sy uiteensetting van die beginsels insake grondwetlike skade in *Maharaj v Attorney General of Trinidad and Tobago (2)* 1979 AC 385 (PC), het die Geheime Raad dit onomwonde gestel dat dit 'n afsonderlike aksiegrond en onafhanklik van 'n deliktuele eis is. In 'n onlangse beslissing, *James v Attorney General of Trinidad and Tobago* 2010 UKPC 23, het die Geheime Raad egter beslis dat 'n eiser wat skade wil eis weens die aantasting van 'n grondwetlik-verskanste reg, skade moet bewys en sodoende een van die wesenlike vereistes vir meeste delikte toegepas. Die artikel voer die argument dat die *James*-benadering nie in lyn is met die *Maharaj*-nalatenskap en die konsep van grondwetlike skade nie.

1 INTRODUCTION

Since the Privy Council laid down the principle in *Maharaj v Attorney General of Trinidad and Tobago (2)*¹ that damages were an “appropriate” relief in a claim for “redress” for breach of fundamental rights entrenched in the Constitution of Trinidad and Tobago 1976,² constitutional damages have been awarded by courts in the Commonwealth in a number of cases.³ In *Maharaj* itself, the appellant was

1 1979 AC 385 (PC) (*Maharaj (2)*).

2 S 14(1) of the 1976 Constitution provides that: “For the removal of doubts it is hereby declared that if any person alleges that any of the provisions of this Chapter has been, is being, or is likely to be contravened in relation to him, then without prejudice to any other action with respect to the same matter which is lawfully available, that person may apply to the High Court for redress by way of originating motion.” It is further provided in s 14(2) that: “The High Court shall have original jurisdiction – (a) to hear and determine any application made by any person in pursuance of subsection (1); and may . . . make such orders, issue such writs and give such directions as it may consider appropriate for the purpose of enforcing, or securing the enforcement of, any of the provisions of this Chapter to the protection of which the person concerned is entitled.”

3 Recent examples of such cases include: Bahamas: *Takitola v Attorney General* 2009 UKPC 12; *Merson v Cartwright* 2005 UKPC 38; Canada: *Vancouver (City) v Ward* 2010 SCC 27 (CanLII); Jamaica: *Brown v Robinson* 2004 UKPC 56; New Zealand: *Minister of Immigration v Udompun* 2005 NZCA 128, 2005 3 NZLR 204 (CA); *Brown v Attorney General* 2006 NZAR; *Percival v Attorney General* 2006 NZAR 215; *R v Williams* 2007 3 NZLR 207 (CA); *McKean v Attorney General* 2007 3 NZLR 819; *Taunoa v Attorney General* 2007 NZSC 70, 2007 5 LRC 680 (NZSC); South Africa: *Zealand v Minister of Justice and*

continued on next page

committed for a term of imprisonment for contempt of court arising from an order made by a High Court Judge in violation of the principles of fundamental justice, and the State was held liable in damages. The damage in this case emanated from the deprivation of the right to personal liberty guaranteed the appellant by the Constitution of Trinidad and Tobago.

In its more recent decision in *Attorney General of Trinidad and Tobago v Ramanoop*,⁴ the Privy Council however added a new dimension to claims for constitutional damages. In that case, a police officer patently abused his office in a naked show of police power in violation of a citizen's right to personal liberty. The question was whether exemplary damages could be awarded by way of redress for contravention of the fundamental rights provisions in the Constitution of the Republic of Trinidad and Tobago. In holding that an "additional award" rather than exemplary damages should be made in this case, the Privy Council made pronouncements which were to influence one way or another, the determination of the claim of the plaintiff in the subsequent case of *James v Attorney General of Trinidad and Tobago*.⁵ Here, two important issues arose in the consideration as to whether damages should be awarded in the event of a constitutional breach. The first is: whether damage must be proved? The second is whether a constitutional breach requires to be marked by an award of damages?

In literally every case decided in the period between *Maharaj* and *James*, the damage involved had constituted physical injury to the person such as assault and battery,⁶ wrongful arrest or false imprisonment or deprivation of right to property. The damage in each case was more or less physical, if not tangible. Any question of proof of damage in those circumstances would therefore have been purely academic. But the situation in *James* is quite different in that there was neither physical⁷ nor material injury, not even financial loss⁸ in terms of loss

Constitutional Development 2008 ZACC 3, 2008 7 BCLR 601, 2008 4 SA 458 (CC); *President of the Republic of South Africa v Modderklip Boerdery (Pty) Ltd* 2005 5 SA 3 (CC), 2005 8 BCLR 786 (CC); St Christopher and Nevis: *Inniss v Attorney General of St Christopher and Nevis* 2008 UKPC 42 (PC); St Lucia: *Fraser v Judicial and Legal Services Commission* 2008 UKPC 25; Trinidad and Tobago: *Attorney General of Trinidad and Tobago v Ramanoop* 2005 UKPC 15, 2006 1 AC 338 (PC); *Subiah v Attorney General of Trinidad and Tobago* 2008 UKPC 47; *Durity v Attorney General (Trinidad and Tobago)* 2008 UKPC 59.

⁴ 2005 UKPC 15, 2006 1 AC 328, 2005 2 WLR 1324 (PC).

⁵ 2010 UKPC 23.

⁶ In the recent New Zealand High Court case of *Falwasser v Attorney General* 2010 NZHC 410 para 132, the plaintiff was assaulted several times with police batons and repeatedly pepper sprayed while in police custody. Stevens J found these to have been excessive and unnecessary and an abuse of power on the part of the police officers involved. As a consequence, it amounted to a failure to treat the plaintiff with humanity and with respect for the inherent dignity of his person and accordingly breached s 23(5) of the New Zealand Bill of Rights Act 1990. He was awarded damages.

⁷ Cf *Attorney General of Trinidad and Tobago v Ramanoop* 2005 UKPC 15, 2006 1 AC 328, 2005 2 WLR 1324 (PC).

⁸ Cf *Shortt v The Commissioner of an Garda Síochána* 2007 IESC 9 where the plaintiff was a victim of disreputable conduct and a shocking abuse of power on the part of two Garda [police] officers. They both engaged in a conspiracy to concoct false evidence against the plaintiff which in turn resulted in perjured Garda evidence being given at his trial for allegedly permitting drugs to be sold in his licensed premises. That perjury procured his conviction by a jury. What followed as a consequence for the plaintiff was a tormenting saga of imprisonment, mental and physical deterioration, estrangement from family, loss of busi-

of earning capacity⁹ or future earning. There was no question of loss of human dignity¹⁰ or reputation.¹¹

In *James*, a police officer filed a constitutional motion seeking declarations and damages for alleged discrimination against him and that he was unfairly treated in relation to the exemption in the English language component of the qualifying examination towards his promotion. Although the trial judge made declarations that the plaintiff had been treated unfairly and discriminated against contrary to section 4(b) and/or 4(d) of the Constitution of Trinidad and Tobago,¹² he briefly and summarily threw out the damages claim because the declaratory relief was sufficient.

2 PROOF OF DAMAGE IN TORT/DELICT

The logic of the law of delict is that if the action of the plaintiff is to seek reparation for the injury sustained in the hands of the defendant, then there should indeed be injury on the basis of which liability can be imposed.¹³ It is thus a requirement of the law of tort,¹⁴ as it is of delict,¹⁵ that the element of damage is essential to establish liability.¹⁶ Not all damage will justify recovery.¹⁷ The

ness, public and professional ignominy and despair. In short, the trial court Judge put it, “the plaintiff was sacrificed in order to assist the career ambitions of a number of members of the Garda Síochána”.

9 *Jorsingh v Attorney General* 1997 3 LRC 333 (T&TCA).

10 *Taunoa v Attorney General* 2007 NZSC 70, 2007 5 LRC 680.

11 *Rees v Crane* 2001 3 LRC 510 (T&TCA).

12 S 4 of the Constitution of Trinidad and Tobago, insofar as is relevant, provides: “It is hereby recognised and declared that in Trinidad and Tobago there have existed and shall continue to exist, without discrimination by reason of race, origin, colour, religion or sex, the following fundamental rights and freedoms, namely: (b) the right of the individual to equality before the law and the protection of the law . . . (d) the right of the individual to equality of treatment from any public authority in the exercise of any functions.”

13 See the discussion in Okpaluba and Osode *Government liability: South Africa and the Commonwealth* (2010) para 8.3.

14 This has been the case in the law of negligence since *Donoghue v Stevenson* [1931] UKHL 3, 1932 AC 562 (HL).

15 *Coetzee v SA Railways and Harbours* 1933 CPD 565 576.

16 It is for this reason that the claims for “wrongful life” or “wrongful birth” have generated much controversy across the common-law world. See eg Australia: *Harriton v Stephens* 2006 HCA 15, 2006 226 CLR 52, 2006 ALR 391; *Cattanach v Melchior* [2003] HCA 38, 2003 77 ALJR 1312; England: *Airedale NHS Trust v Bland* 1992 UKHL 5, 1993 AC 789; *Re B (adult, refusal of medical treatment)* 2002 2 All ER 449; *McKay v Essex Area Health Authority* 1982 2 All ER 771 (CA); South Africa: *Stewart v Botha* 2007 6 SA 247 (C); 2008 6 SA 310 (SCA); *Friedman v Glicksman* 1996 1 SA 1134 (W); *Administrator of Natal v Edouard* 1990 3 SA 581 (A). See also Teff “The action for ‘wrongful life’ in England and the United States” 1985 *ICLQ* 423; Luntz *Assessment of damages for personal injury and death* (2002) 641 para 11.8.8; Luntz and Hambly *Torts: Cases and commentary* (2002) 433 para 7.2.19; Hayes and Hayes “Permanent disability and common law rights to compensation” 1982 *ALJ* 643; Fleming *The law of torts* (1998) 184–185.

17 On the question as to what damage is capable of being protected by the law, the case of *Du Plessis v RAF* 2004 1 SA 359 (SCA) is instructive. The plaintiff was in a same-sex permanent life relationship with the deceased until he was killed in a motor vehicle accident. Having established that the deceased owed the plaintiff a contractual duty of support, the next question the court had to determine was whether the right of the plaintiff to such support was worthy of protection by way of an action against the defendant. After considering the prevailing *boni mores*, translated into the law of delict’s “legal convictions of the community” viewed from the constitutional tenets, rights and values, including the right to

continued on next page

damage or injury generally recoverable is damage to person, property or pure economic loss.¹⁸ While economic loss may concern the loss of a property right, financial benefit, business interest, loss of profit, means of livelihood or loss of earning capacity, physical injury may relate to injury to the human person, personal liberty, distress and inconvenience, loss of reputation, or psychiatric injury.¹⁹ Personal injury also includes psychological injury.²⁰ The courts have thus refrained from drawing a distinction between physical and mental injury for they find it somewhat artificial to do so.²¹

Having shown that the plaintiff suffered damage, he/she must go further and show that “but for” the defendant’s conduct the harm would not have occurred.²² Depending on whether that question was answered in the affirmative, the next stage is to show that the wrongful act was linked sufficiently closely or directly

equal protection of the law and the right to human dignity as elucidated in the now well developed jurisprudence of the Constitutional Court, the Supreme Court of Appeal held that the plaintiff, as a same-sex partner of the deceased in a permanent life relationship similar in all other respects to marriage, in which the deceased had undertaken a contractual duty of support to him, was entitled to claim damages from the defendant for the loss of that support. Similarly, in *Grütter v Lombard* 2007 4 SA 89 (SCA), where the question was whether the features of a person’s identity constitute interests capable of legal protection, it was held that they are capable of such protection and that the *actio injuriarum* was capable of protecting a person against unauthorised publication of his name. However, not every intrusion upon those protectable rights of personality will necessarily constitute an *injuria*. Whether a particular act constitutes a wrongful (or unlawful) violation, and thus an *injuria*, must necessarily be determined by considerations of legal policy.

18 For the controversy surrounding claims for pure economic loss and the boundaries of pure economic loss and psychological harm, see Okpaluba and Osode *Government liability: South Africa and the Commonwealth* (2010) para 6.3. See also Klar *Tort law* (2003) 201–214; Weir *Tort law* (2002) 44–51.

19 *Hill v Hamilton-Wentworth Regional Police Services Board* 2007 SCC 41, 2007 285 DLR (4th) 620 para 90. In paras 91–92, the majority held that: “It was not disputed that imprisonment resulting from a wrongful conviction constitutes personal injury to the person imprisoned. Indeed, other forms of compensable damage without imprisonment may suffice; a claimant’s life could be ruined by an incompetent investigation that never results in imprisonment or an unreasonable investigation that does not lead to criminal proceedings. Wrongful deprivation of liberty has been recognised as actionable for centuries and is clearly one of the possible forms of compensable damage that may arise from a negligent investigation. There may be others. On the other hand, lawful pains and penalties imposed on a guilty person do not constitute compensable loss. It is important as a matter of policy that recovery under the tort of negligent investigation should only be allowed for pains and penalties that are wrongfully imposed. The police must be allowed to investigate and apprehend suspects and should not be penalised for doing so under the tort of negligent investigation unless the treatment imposed on a suspect results from a negligent investigation and causes compensable damage that would not have occurred but for the police negligent conduct. The claimant bears the burden of proving that the consequences of the police conduct relied upon as damages are wrongful in this sense if they are to recover. Otherwise, punishment may be no more than a criminal’s just deserts – in a word, justice.”

20 *Mustapha v Culligan of Canada Ltd* 2008 SCC 27 (CanLII) para 9 per Lachlin CJC. See also *Hinz v Berry* 1970 2 QB 40 (CA) 42; *Page v Smith* 1996 1 AC 155 (HL) 189; Linden and Feldthusen *Canadian tort law* (2006) 425–427.

21 Per Lord Lloyd in *Page v Smith* 1996 1 AC 155 (HL) 188.

22 *Mustapha v Culligan of Canada Ltd* 2008 SCC 27 (CanLII) para 13.

to the loss.²³ The law is that where the injury is too unrelated or remotely related to the wrongful conduct, the defendant will not be held liable.²⁴

In South Africa, the Supreme Court of Appeal took the opportunity presented by the case of *Jowell v Bramwell-Jones*²⁵ to restate the law. First, the element of damage or loss is fundamental to the *Aquilian* action and the right of action is incomplete until damage is caused to the plaintiff by reason of the defendant's wrongful conduct.²⁶ Secondly, this requirement applies no less to claims arising from pure economic loss than it does to claims arising from bodily injury or damage to property.²⁷ Thirdly, whether a plaintiff has suffered damage or not is a fact which, like any other element of his cause of action, must be established on a balance of probabilities. Finally, once the damage or loss is established, a court will do its best to quantify that loss even if this involves a degree of guesswork.²⁸

What emerges from the requirement of damage is that there must be proof on a balance of probabilities such that where no evidence of loss is shown, no cause of action accrues.²⁹ The requirement of damage also has implications for quantification; the damage must be capable of being quantified. Although the actual or material damage as a cause of the injury and quantification of that damage emanate from the same source, a distinction exists between the two. Both have distinct roles to play in the law of delict.³⁰

3 THE COURT OF APPEAL IN *JAMES*

At the Court of Appeal, it was contended that the declaratory relief was not sufficient recompense for the wrong suffered by the appellant. It was argued further that the giving effect to and protecting and securing the enforcement of the fundamental rights and freedoms guaranteed under section 4(b) and (d) of the Constitution required the award of damages including aggravated and/or exemplary damages. The aggravation of the appellant's injury was said to have

23 *mCubed International (Pty) Ltd v Singer NNO* 2009 4 SA 471 (SCA) para 22; *Fourway Haulage SA (Pty) Ltd v SA National Roads Agency Ltd* 2009 2 SA 150 (SCA) para 31; *South Australia Asset Management Corporation v York Montague Ltd* 1997 AC 191, 1996 3 All ER 365 (HL) 371j.

24 See *Overseas Tankship (UK) Ltd v Morts Dock & Engineering Co* 1967 AC 617 (PC) 643 (*The Wagon Mound No 2*) reformulating its earlier views in *The Wagon Mound (No 1)* 1961 AC 388 (PC) 424. See also Linden and Feldthusen *Canadian tort law* (2006) 360.

25 2000 3 SA 274 (SCA) 286G–I.

26 *Oslo Land Co Ltd v Union Government* 1938 AD 584 590; *Evins v Shield Insurance Co Ltd* 1980 2 SA 814 (A) 838H–839C.

27 *Siman & Co (Pty) Ltd v Barclays National Bank Ltd* 1984 2 SA 888 (A) 911B–D.

28 *Turkestra Ltd v Richards* 1926 TPD 276 282–3.

29 The judgment of Jones J in *Reivelo Leppa Trust v Kritzinger* 2007 3 All SA 794 (SE) 798–99 captures all the issues involved here. The plaintiff claimed damages arising from fraudulent or negligent misrepresentation or non-disclosure in connection with an agreement of purchase and sale between the parties. In the exception raised by the defendant, it was contended that the plaintiff had failed to allege a cause of action, or a complete cause of action, for the relief against the trust, and that it had not alleged a legally valid basis for claiming costs against the second defendant. The trial judge upheld the exception and held that the plaintiff had done no more than raise a possibility that he may suffer some form of undetermined harm, which he may be able to quantify at a later stage. However, that is insufficient for a cause of action. There were therefore no proper averments to sustain a cause of action.

30 Midgley and Van der Walt "Delict" in Joubert (ed) 8 *LAWSA* 266.

stemmed from the manner in which the State had conducted the case against the appellant in insisting that he had not suffered discrimination and that he was not entitled to the exemption as he had claimed.

The Court of Appeal of Trinidad and Tobago dismissed the appeal. Kangaloo JA who delivered the judgment of the court held that insofar as the trial judge decided that damages should not be awarded without offering explanations, it was for the appellate court to consider the matter afresh and arrive at a view as to how the discretion to award damages should be exercised especially as the facts were not in dispute. Relying on the language of section 14(2) of the Constitution and the Privy Council decision in *Suratt v Attorney General of Trinidad and Tobago*,³¹ Kangaloo JA held that damages were “a subset of the discretionary relief of the court in granting relief”. Further offering a significant qualification to this statement, he held that before damages could be awarded, it was “obvious” that they must relate to some loss suffered by the applicant.

Again, referring to *Ramanoop*,³² the Justice of Appeal held that it was beyond doubt that an applicant must demonstrate that as a result of the breach of his constitutional right he has suffered damage, as much as it remains within the discretion of the court as to whether or not to make a monetary award. In the absence of any evidence that the breach of the appellant’s right caused him some damage, the court held that the appellant had suffered no monetary loss. A mere mention that the appellant suffered distress and inconvenience was not enough to move the court to exercise its discretion to award him damages. Kangaloo JA reasoned:

“[I]t must first be shown that there has been damage suffered as a result of the breach of the constitutional right before the court can exercise its discretion to award damages in the nature of compensatory damages to be assessed. If there is damage shown, the second stage of the award is not available as a matter of course. It is only if some damage has been shown that the court can exercise its discretion whether or not to award compensatory damages. The practice has developed in constitutional matters in this jurisdiction of having a separate hearing for the

31 2008 UKPC 38 (PC) para 12. In this case damages were claimed because of the failure on the part of the government to implement a law. In holding that there was no sound basis for an award of damages, vindictory or otherwise, Lord Brown held: “In the present case, indeed, the Board conclude that no compensatory award is required either. By the same token that the House of Lords in *R (Greenfield) v Secretary of State for the Home Department* 2005 UKHL 14, 2005 1 WLR 673 held that United Kingdom courts, exercising the power to award damages or order compensation given to them under s 8 of the Human Rights Act 1998 should, consistently with the ECtHR exercising its A 41 jurisdiction under the European Convention on Human Rights, treat the finding of the violation, certainly in a 6 cases, as in itself affording just satisfaction to the injured party rather than speculate on what the outcome of the proceedings would have been but for the violation, and generally should award damages only where satisfied that loss had in fact been caused by the violation, so too their Lordships believe that the Board’s declarations already made on 15 October 2007 provide adequate and appropriate redress for the Appellants here. Obviously the analogy with the application to due process violations of s 8 of the Human Rights Act (and a 41 of the Convention) is not an exact one. But what the Appellants have been deprived of here assuming always in their favour that their s 4(b) constitutional right has been breached is the opportunity in recent years to bring discrimination complaints against private persons. What the upshot of any such complaints would have been is in the highest degree speculative.”

32 *Attorney General of Trinidad and Tobago v Ramanoop* 2005 UKPC 15, 2006 1 AC 338 (PC).

assessment of the damages, but it cannot be overemphasized that this is after there is evidence of the damage. In the instant case there is no evidence of damage suffered as a result of the breaches for which the appellant can be compensated.³³

Although the Court of Appeal considered that the fact that a compensatory award had not been made did not necessarily preclude the making of an additional award as enunciated by Lord Nicholls in *Attorney General of Trinidad and Tobago v Ramanoop*,³⁴ it found nothing in the case that showed that public outrage was excited. There was no doubt that the constitutional right breached was important but there was no need for deterrence nor was there *mala fides* on the part of the Police Service. Additional award therefore did not arise.

4 THE JUDGMENT OF THE PRIVY COUNCIL

Before the Privy Council, the appellant contended that the Court of Appeal erred in imposing as a requirement for access to monetary compensation that there be damage over and beyond that which naturally and inevitably flows from being the victim of discrimination and unequal treatment in the constitutional context. Distress and anxiety on the part of the appellant at the wrongful denial of his exemption must be presumed. The denial of monetary compensation in this case would significantly devalue the appellant's constitutional right and make light of its serious breach by the authorities. The appellant argued that he was entitled to additional or vindicatory damages to cater for the factors identified in *Ramanoop*. It was however submitted for the respondent that the Court of Appeal had properly exercised its discretion in refusing to make either a compensatory or an additional award in this case. Accordingly, compensation for distress and inconvenience cannot be presumed; it must be proved. The declaration that the appellant's right has been violated was sufficient relief and the appellant could not receive any extra benefit at the discretion of the court.

It was held that the right which the appellant has asserted is not one which, absent the erroneous grant of the exemption to other officers, would have been available to him. He has been treated unequally only because others have been treated better than he (and better than they ought to have been) due to an administrative error. If the rules had been properly applied to all, neither the appellant nor those to whom he has compared himself in order to demonstrate unequal treatment, would have received the exemption.³⁵ Lord Kerr further held that the

33 Quoted by Lord Kerr 2010 UKPC 23 para 13.

34 2005 UKPC 15, 2006 1 AC 338 (PC) para 19.

35 2010 UKPC 23 para 22. It was held by Lord Carswell in *Bhagwandeem v Attorney General of Trinidad and Tobago* 2004 UKPC 21 para 18 that for a claimant who alleges inequality of treatment or its synonym discrimination must ordinarily establish that he has been or would be treated differently from some other similarly circumstanced person or persons, described by Lord Hutton in *Shamoon v Chief Constable of the Royal Ulster Constabulary* [2003] 2 All ER 26 para 71 as actual or hypothetical comparators. The phrase which is common to the anti-discrimination provisions in the legislation of the United Kingdom is that the comparison must be such that the relevant circumstances in the one case are the same, or not materially different, in the other. The appellant in *Bhagwandeem* relied on an actual comparator, sergeant Fitzgerald George, who was promoted some seven and three quarter months after his reinstatement following a long period of suspension. That interval of time is the only real similarity between their cases. In his affidavit sworn on 28 December 1999 the Commissioner compared the history and record of the appellant and sergeant George in some detail, which it was not necessary to set out in detail. Their lordships were

continued on next page

enforcement of the protective provisions may require more than mere recognition that a violation of those provisions has occurred. In effect, “when exercising this constitutional jurisdiction the court is concerned to uphold, or vindicate, the constitutional right which has been contravened”.³⁶ The constitutional dimension adds an extra ingredient. The violated right requires emphatic vindication. For that reason, careful consideration must be given to the nature of the breach; the circumstances in which it occurred; and the need to send a clear message that it should not be repeated.

Frequently, this will lead to the conclusion that something beyond a mere declaration that there has been a violation will be necessary. This is not inevitably so, however. Nor is it even the case that it will be required in all but exceptional circumstances. Close attention to the facts of each case is required in order to decide on what is required to meet the need for vindication of the constitutional right which is at stake.³⁷

Lord Kerr held that Kangaloo JA had in mind some type of damage beyond that of being the victim of a violation of an applicant’s constitutional rights before the question of compensation could arise. The Justice of Appeal founded this conclusion on the third sentence of Lord Nicholls’ speech to the effect that: “*If the person wronged has suffered damage, the court may award him compensation.*” However, His Lordship did not purport to propound a general rule that some form of extra or superimposed damage was required before monetary compensation could be considered. In the context of constitutional violation, compensation can be seen to perform two functions. It may, of course, provide redress for the *in personam* damage suffered. But it may also be an essential part of the vindication of the constitutional right. This has a broader concept than compensation for the personal wrong. Although it is clear that compensation in this area should not include an exemplary or punitive element³⁸ – part of the function of an award of compensation in this setting is to mark the fact that a *constitutional* breach has occurred. Lord Nicholls was therefore observing that when that extra ingredient is required, one of the forms that it may take is monetary compensation when the person who has been the victim of the breach of the constitutional protection has suffered damage. He was not suggesting that some specific type of damage suffered by the victim of the constitutional breach was necessary before the question of monetary compensation could be considered.³⁹

In any event, the very fact of discrimination having occurred can inflict damage on those who have been discriminated against. The sense of having been wronged, the uncertainty over one’s status as a consequence of the discriminatory conduct and the distress associated with having to resort to litigation in order to have the discrimination exposed and corrected can all be recognised as damage,

satisfied on this evidence, which was not controverted, that the Court of Appeal was correct in rejecting sergeant George as a true comparator. Since the appellant had to concede that none of the other officers to whom he referred in his affidavit sworn on 5 August 1999 as having been promoted following periods of suspension was similarly circumstanced to himself, he was left without a foundation for his claim of unequal treatment.

36 2005 UKPC 15, 2006 1 AC 338 (PC) para 18 per Lord Nicholls.

37 2010 UKPC 23 para 24.

38 Per Lord Nicholls in *Ramanoop* 2005 UKPC 15, 2006 1 AC 338 (PC) para 19.

39 2010 UKPC 23 para 26.

perhaps not in the conventional personal injury sense, but damage nonetheless.⁴⁰ An injury suffered as a result of discrimination is no less real because it does not possess tangible physical or financial consequences. And the difficulty in assessing the amount of compensation for that type of injury should not deter a court from recognising its potential to be compensated.⁴¹

It was further held that although it may be easy to prove in terms of persuasion that the anger, distress and affront caused by the act of discrimination have injured the applicant's feelings, compensation for injury to feeling is not automatic;⁴² it is compensatory.⁴³ Their Lordships therefore concluded that, although the Court of Appeal was wrong in its view that there needed to be some form of additional damage (beyond being the victim of discrimination) before the question of damages could be considered, it was unquestionably right in its decision that there was no acceptable evidence that the appellant had suffered an injury to his feelings or particular distress.⁴⁴ To put the matter differently, although, by

40 Per Lord Kerr 2010 UKPC 23 para 27.

41 *Ibid.* On this see *Vento v Chief Constable of West Yorkshire Police* 2002 EWCA Civ 1871, 2003 ICR 318 331, 2003 IRLR 102 (CA) paras 50 and 51 where, dealing with the quantification of damages in a sex discrimination case, which like racial discrimination, presents similar problems of quantification, Mummery LJ observed that the assessment of compensation for an injury or loss, which is neither physical nor financial, presents special problems for the judicial process, which aims to produce results objectively justified by evidence, reason and precedent. Subjective feelings of upset, frustration, worry, anxiety, mental distress, fear, grief, anguish, humiliation, unhappiness, stress, depression and so on and the degree of their intensity are incapable of objective proof or of measurement in monetary terms. Translating hurt feelings into hard currency is bound to be an artificial exercise. This is because there is no medium of exchange or market for non-pecuniary losses and their monetary evaluation "is a philosophical and policy exercise more than a legal or logical one. The award must be fair and reasonable, fairness being gauged by earlier decisions; but the award must also of necessity be arbitrary or conventional. No money can provide true restitution". See also *Heil v Rankin* 2001 QB 272 292 para 16; *Andrews v Grand & Toy Alberta Ltd* 1978 83 DLR (3d) 452 475–476 per Dickson J.

42 *Ministry of Defence v Cannock* 1994 ICR 918 954 (EAT).

43 *HM Prison Service v Johnson Smith* 1997 ICR 275 (EAT) 283B per Smith J. See also *Gbaja-Bianila v DHL International (UK) Ltd* 2001 ICR 730 (EAT); *ICTS (UK) Ltd v Tchoula* 2000 IRLR 643; *HM Prison Service v Salmon* 2002 IRLR 425 (EAT).

44 Cf in *Rees v Crane* 2001 3 LRC 510 521–524 (Trinidad and Tobago CA) where damages for distress and inconvenience were awarded on account of the manner a judge of the High Court was treated by the authorities. While observing that distress and inconvenience were abstract elements difficult to express in monetary terms, Hamel-Smith JA spoke of the essential factors in this case in these words: "In a nutshell, he complained that the breach of his right to be heard at the appropriate time left him in shambles. He was first shocked by his not being rostered to sit in court and then humiliated by learning of his suspension from office over the national television. He was further insulted and humiliated by having been served with the instrument of suspension by a policeman on a street corner in the full gaze of the public. The action on the part of the commission brought him untold anxiety and grief. It made him appear to be guilty of some gross misconduct without being given particulars of the allegations and without being given an opportunity to deal with them at any stage. This caused him to suffer from sleepless nights over a prolonged period, a condition that affects him to this day. His reputation as a judge in his view has been severely tarnished, if not destroyed, and this has caused him severe distress. He had to suffer the indignity of loss of office of a judge for some three years and more while his action was pending in court and the feeling of being shunned by society as a result. Further, the action of the commission caused him loss of opportunity of promotion to the Court of Appeal and loss of prospects of employment and professional income."

dint of the fact that the appellant was found to have been the victim of discrimination, the appellant was entitled to have access to the court's judgment whether to award damages, insofar as monetary compensation depended on an assessment of the impact on him, there was simply no basis on which it could be made.⁴⁵

5 MUST A BREACH OF RIGHT BE ACCOMPANIED BY AN AWARD OF DAMAGES?

In neither *Suratt* nor *Ramanoop* did the Privy Council decide that a constitutional breach must necessarily require that an award of damages must follow. This is so because damages are but one of a host of remedies from which a court could choose to ventilate a fundamental right infraction.⁴⁶ Granted that the appellant in this case has established that he was a victim of discrimination compared with what the appellants in *Suratt* lost, that is, the opportunity to bring proceedings against individuals and there was no guarantee that, had that opportunity been available, they would have succeeded in any claim. Yet, that alone does not guarantee that entitlement to compensation was automatic as contended by the appellant in the present case. This proposition, contended the appellant, is supported by the opinion of Lord Nicholls in *Ramanoop*⁴⁷ already quoted above. Responding, Lord Kerr observed:

"I do not understand Lord Nicholls in this passage to be suggesting that, in principle, compensation should normally be awarded. What, as it seems to me, he was at pains to point out was that a violation of someone's constitutional rights will commonly call for something more than a mere statement to that effect. This is required in order to reflect the importance of the constitutional right and the need for it to be respected by the state authorities. A risk of the devaluation of such rights would obviously arise if the state could expect that the most significant sanction for their being flouted was a declaration that they had been breached. It is, of course, significant that in *Ramanoop* there was no dispute as to whether the respondent was entitled to damages. The question in that case was whether the state should be required to pay an additional amount of damages in order to reflect the sense of public outrage that the breach of the constitutional right had occasioned. Moreover, the case was one which clearly called for a compensatory award (as well as the additional award). The respondent had been assaulted over a prolonged period by a police officer. The latter's behaviour had been described by Lord Nicholls as 'quite appalling'. The need for a compensatory award and, indeed, an additional award of damages was, in that case, quite plain."⁴⁸

It was held that to treat entitlement for compensation as automatic where violation of a constitutional right has occurred would undermine the discretion that is invested in the court by section 14 of the Constitution of Trinidad and Tobago 1976. Again, it would run directly counter to the jurisprudence in this field of law. For instance, Lord Hope made it clear in *Inniss v Attorney General of St Christopher and Nevis*⁴⁹ that it is important to consider whether a declaration that

45 Per Lord Kerr, 2010 UKPC 23 para 32.

46 See Okpaluba "Of forging new tools' and 'shaping innovate remedies': Unconstitutionality of legislation infringing fundamental rights arising from legislative omissions in the new South Africa" 2001 *Stell LR* 462 464–469.

47 2005 UKPC 15, 2006 1 AC 338 (PC) para 18.

48 2010 UKPC 23 para 35.

49 2008 UKPC 42 para 21.

there was a contravention of section 83(3) of the Constitution of St Christopher and Nevis would be sufficient relief and whether such a remedy would serve to vindicate the constitutional right infringed. In some cases a declaration on its own may achieve all that is needed to vindicate the right in the circumstances of the case. For instance, in a situation where the contravention has not yet had any significant effect or major impact on the claimant, a declaration is more likely than not to suffice. The present case is a notable illustration. Even though it was found that the appellant had been discriminated against, he did not appear to have been affected in any material way whatever. Despite having been denied the exemption, he was in fact promoted and has continued in the rank of acting sergeant even after the error was detected. He has now secured an exemption through the legal proceedings that he instituted. His case does not fit the circumstances envisaged in *Inmiss*. The appellant has not suffered any injury nor will he suffer any, a declaratory judgment having already been rendered.⁵⁰

The appeal against the refusal to award the appellant compensatory damages was therefore dismissed.⁵¹ It was similarly held that this was plainly not one in which an additional award could be made. It was not in doubt that the constitutional right infringed was an important right but the breach was of an eccentric kind, consisting as it did of a failure to include the appellant in a departure from the rules governing a promotion which he had in the event obtained. The possibility of further breaches of similar character arising was particularly remote. It follows therefore that no case was made that the considerations that underlie such an award were present. The appeal was also dismissed on this ground.⁵²

6 ANALOGY WITH *WATKINS*

James compares with two earlier judgments of the House of Lords emanating from the same juridical school of thought. The first is *Wainwright v Home Office*⁵³ the judgment of which was handed down before the United Kingdom

50 2010 UKPC 23 para 37.

51 *Ibid* para 39.

52 *Ibid* para 42.

53 2004 2 AC 406 para 34. It was held that there was no protection in English law for the invasion of the right to privacy even though that right is guaranteed by Art 8 of the European Convention as incorporated into the Human Rights Act 1998. Accordingly, the prison authorities in *Wainwright* were not liable in damages for a tort that was not known to the law. Lord Hoffman held that the coming into effect of the Human Rights Act weakened the argument for saying that a general tort of invasion of privacy is needed to fill the gaps in the existing remedies. According to him: "Sections 6 and 7 of the Act are in themselves substantial gap fillers; if it is indeed the case that a person's rights under article 8 have been infringed by a public authority, he will have a statutory remedy." In rejecting the view of the Court of Appeal that the Wainwrights had a strong case for relief under s 7 of the Act had it been in force, Lord Hoffmann held (paras 51 and 52): "Speaking for myself, I am not so sure. Although article 8 guarantees a right to privacy, I do not think that it treats that right as having been invaded and requiring a remedy in damages, irrespective of whether the defendant acted intentionally, negligently or accidentally. It is one thing to wander carelessly into the wrong hotel bedroom and another to hide in the wardrobe to take photographs. Article 8 may justify a monetary remedy for an intentional invasion of privacy by a public authority, even if no damage is suffered other than distress for which damages are not ordinarily recoverable. It does not follow that a merely negligent act should, contrary to the general principle, give rise to a claim for damages for distress because it affects privacy rather than some other interest like bodily safety . . . a finding that

continued on next page

Human Rights Act 1998 came into force and, the second, *Watkins v Secretary of State for the Home Department and Others*⁵⁴ which was decided when the Act was already in operation but which, for its closer affinity with the present case, is discussed in this context.

Watkins brought to the fore the contradictions in the approaches of the English Court of Appeal and the House of Lords on the issue whether what the court was called upon to enforce could properly be styled “constitutional”, “basic” or “fundamental” right. It is also clear that the judicial activism discernible from the New Zealand Court of Appeal’s constitutional approach to the interpretation of the provisions of the New Zealand Bill of Rights Act 1990 since *Baigent*,⁵⁵ had reverberated in the English Court of Appeal.⁵⁶ The latter had, in *Watkins*, developed a jurisprudential approach that embraces “constitutional”, “basic” and “fundamental” rights vocabulary in dealing with the types of issues with which a number of tort claims were adjudicated. On the contrary, the House of Lords do not at all pretend to be interpreting an instrument equivalent to a constitution when it is dealing with any aspect of the Human Rights Act 1998. What is perceptible from the attitude of their Lordships is that unless a right was protected at common law, there would be no liability even for a breach of a Convention right.⁵⁷

Watkins was an action for misfeasance in public office where the plaintiff, a serving prisoner, sought damages for the unlawful interference by prison officials who had opened his correspondence with his legal advisers contrary to Prison Rules. The question turned on whether an action for misfeasance in public office can be successfully maintained without special proof of damage. In the language of the English law of tort, is the tort of misfeasance in public office, like trespass, actionable *per se*? In allowing the action, the Court of Appeal held that where the wrongful and malicious act of a public officer interfered with a constitutional right of the claimant, he had a cause of action for misfeasance in public office without the need to prove special damage. The Court of Appeal had likened the interference with the prisoner’s correspondence with his legal adviser contrary to Prison Rules to the infringement of the right to vote in light of the

there was a breach of article 8 will only demonstrate that there was a gap in the English remedies for invasion of privacy which has been filled by sections 6 and 7 of the 1998 Act. It does not require that the courts should provide an alternative remedy which distorts the principles of the common law.” Cf *Hicks v Chief Constable of South Yorkshire Police* 1992 2 All ER 65.

54 2004 EWCA Civ 966, 2005 2 WLR 1538 (CA).

55 *Simpson v Attorney General* [1994] 3 NZLR 667 (NZCA) (*Baigent’s Case*). In addition to the more recent cases cited in fn 3, see also *Dunlea v Attorney General* [2000] 3 NZLR 136; *Attorney General v Hewitt* [2001] NZAR 148 (HC); *Binstead v Northern Region Domestic Violence (Programmes) Approval Panel* [2002] NZAR 865.

56 *Watkins v Secretary of State for the Home Department* 2004 EWCA Civ 966, 2005 2 WLR 1538 (CA). See also *R v Secretary of State for the Home Department, Ex parte Leech* 1994 QB 198 201A–B where Steyn LJ (as he then was) spoke of the “basic right of access to the court” as ranking among “constitutional rights” “even in our unwritten constitution”. And, in *Leech v Secretary for Scotland* 1992 SC 89 98, Lord McCluskey spoke of “a basic right of access to the courts”. See also *R v Lord Chancellor, Ex parte Witham* 1998 QB 575 585; per Lord Steyn, *R v Secretary of State for the Home Department, Ex parte Simms* 2000 2 AC 115 130D–E, and Lord Hoffman 131E–G.

57 See *Wainwright v Home Office* 2004 2 AC 406 para 34 discussed *infra*.

judgment of Lord Chief Justice Holt in *Ashby v White*.⁵⁸ For Brooke LJ, the requirement of proof of damage was not necessary where the defendants had infringed “a right which may be identified as a constitutional right”.⁵⁹ Laws LJ held that this was an instance where the public officer’s unlawful conduct has interfered with a constitutional right.⁶⁰

The House of Lords held that proof of special damage had consistently either been expressly recognised or assumed as an essential ingredient of the tort of misfeasance in public office, and a rule which had represented the law for over 300 years should not be disturbed without compelling reasons for doing so. In holding that the tort of misfeasance in public office “is never actionable without proof of material damage”,⁶¹ Lord Bingham emphasised that there is no existing right to damages where misfeasance in public office has caused no material damage to the victim, “and if the evidence showed an egregious and deliberate abuse of power by a public officer one would expect the Strasbourg court to award compensation for non-pecuniary loss even though its practice is not to award exemplary damages.”⁶²

With the concurrence of the other Law Lords, Lord Bingham deprecated the Court of Appeal’s use of the expressions “constitutional”, “basic” or “fundamental” in describing certain rights such as, in the present instance, the right to protection of the confidentiality of the claimant’s legal correspondence. Accordingly, he saw “scant warrant for importing this jurisprudence into the definition of the tort of misfeasance in public office.”⁶³ Among the reasons advanced by Lord Bingham and accepted by the House for refusing to endorse the “novel steps” taken by the Court of Appeal was that to treat the character of the right invaded as to whether it was “constitutional” or not, as determinative of whether material damage needed to be proved would, in the absence of a codified constitution, be imprecise and controversial. It would open the door to argument as to whether other less obviously “fundamental” or “basic” rights were sufficiently close or analogous to be treated for damage purposes in the same way. Another reason for resisting the award of damages in this case is that the plaintiff was not seeking to be compensated but to punish the defendant, that being the function of exemplary damages.⁶⁴

Lord Rodger held that the novel use of the concept of a “constitutional right” or “a right of this level of importance” to create a type of misfeasance in public office which is actionable *per se* would, for it to be workable, “have to be possible to identify fairly readily what were to count as ‘constitutional rights’ for this purpose in a country without a written constitution.”⁶⁵ In his view, “it might be worth trying to deploy the concept of constitutional rights in the law of tort if it represented a way forward which fitted the present state of the law. But it does not.”⁶⁶ Lord Rodger held that breach of a fundamental human or constitutional

58 1703 1 Smith’s LC 253.

59 2005 QB 883 paras 48 and 49.

60 *Ibid* para 67.

61 2006 UKPC 17, 2006 2 WLR 807 (HL) 817–818 para 27.

62 *Ibid* 817–818 para 26 citing *BB v United Kingdom* 2004 39 EHRR 635 para 36.

63 *Ibid* 817–818 para 25.

64 *Ibid* 817–818 para 26.

65 *Ibid* 826 para 58.

66 *Ibid* 829 para 64.

right would, in all probability, found a claim under section 7 of the Human Rights Act 1998, as it would in this case where the violation occurred after the Act came into force. However, his lordship was of the opinion that the facts of *D v East Berkshire Community Health NHS Trust*⁶⁷ were such that the application of familiar principles supported recognition of a remedy in tort, not a case like the present where the application of settled principle points strongly against one.⁶⁸ Referring to the references to “constitutional rights”, Lord Rodger held that they were made in situations before the Act imported the European Convention into English law; thus, the judges were, more or less explicitly, “looking for a means of incorporation *avant la lettre*, of having the common law supply the benefits of incorporation without incorporation.”⁶⁹ According to him:

“Now that the Human Rights Act 1998 is in place, such heroic efforts are unnecessary: the Convention rights form part of our law and provide a rough equivalent of a written code of constitutional rights, albeit not one tailor-made for this country. In general, at least, where the matter is not already covered by the common law but falls within the scope of a Convention right, a claimant can be expected to invoke his remedy under the Human Rights Act rather than to seek to fashion a new common law right.”⁷⁰

The reasoning in *Watkins* clearly removed the claim from being considered as a breach of a constitutional right since in the view of their Lordships no such right exists in the absence of a written constitution. The obvious implication of that approach is that the plaintiff would go the whole distance in establishing liability in tort. As in the present case, the plaintiff must show:⁷¹ first, that the officer in question was a public officer performing a public duty; second, that he acted in bad faith, maliciously or recklessly; third, that his action caused a material or special damage which was not too remote. If, on the other hand, the matter comes to court as a constitutional motion, none of these questions need be asked or answered. As the Indian Supreme Court puts it, the liability of the state in this regard is one of “strict liability”.⁷² No more than a breach of a fundamental right is required to ground liability. Thus, to proceed in accordance with the *Maharaj* injunction would by-pass all these hurdles placed by the law of negligence or misfeasance in public office.

7 COMPARING *KARAGOZLU* AND *ZEALAND*

It has already been observed that the House of Lords threw out the claim based on misfeasance in public office in *Watkins v Secretary of State for the Home Department and Others*⁷³ because the claimant was unable to prove that he

67 2005 UKHL 23, 2005 2 AC 373 (HL) para 50.

68 2006 UKPC 17, 2006 2 WLR 807 (HL) 818 para 26.

69 *Ibid* 829 para 64.

70 *Ibid* para 64. See also per Lord Hoffman in *Wainwright v Home Office* 2004 2 AC 406 para 33. For Lord Walker (para 73), if the facts of the present case occurred again, a prisoner would have a clear claim under ss 6, 7 and 8 of the Act by reference to both aa 6 and 8 of the Act.

71 See eg *Northern Territory of Australia v Mengel* 1995 185 CLR 307 (HCA); *Garrett v Attorney General* 1997 2 NZLR 332 (CA); *Three Rivers DC v Bank of England* 2000 3 All ER 1 (HL), 2002 2 WLR 1220; *Odhavji Estate v Woodhouse* 2004 233 DLR (4th) 193 (SCC).

72 Per Verma J in *Nilabati Behera v State of Orissa* 1994 2 LRC 99 (India SC) 105g/h.

73 2006 2 WLR 807 (HL).

suffered special damage in consequence of the public officer's wrongful conduct. In other words, in the absence of any special damage caused to the claimant by the actions of the prison officers, his claim for damages failed. Unlike the *Watkins* case, the subsequent decision of the English Court of Appeal in *Karagozlu v Commissioner of Police of the Metropolis*⁷⁴ concerned a serving prisoner's right to personal liberty.

The appellant in *Karagozlu* was a category D prisoner in an open prison who was transferred to a closed category B prison. The reason for his transfer was said to be for his safety which the claimant contended was based on false information provided by a police officer and maliciously passed on to the prison authorities, ostensibly to harm him. He commenced an action against the police for damages for misfeasance in public office. And the question was whether, on the assumed facts, the claimant had suffered damage for the purposes of the tort of misfeasance in public office. The Court of Appeal had no doubt that in the light of *Watkins*, loss or damage was an essential element of the tort of misfeasance, but the question it had to answer was: what amounts to damage for this purpose? The submission was that in *Watkins*, the House of Lords was not called upon as the Court of Appeal in the instant case, to consider a situation where a person loses his freedom as a result of a dishonest abuse of power by a public servant such as the police officer in this case. The plaintiff had, therefore, suffered damage sufficient to entitle him to recover general damages for that loss of freedom.⁷⁵

The Court of Appeal took the analogy of the torts of false imprisonment and malicious prosecution where a person who is unlawfully detained and loses his freedom is thereby entitled to general damages.⁷⁶ Indeed, in *Thompson v Commissioner of Police of the Metropolis*,⁷⁷ the Court of Appeal went as far as laying down the guidelines for the directions to the jury in false imprisonment and malicious prosecution claims in respect of basic damages, aggravated damages and exemplary damages.⁷⁸ The crucial issue was the loss of liberty, and the question turned on whether, on the assumption that those allegations of fact were true, the plaintiff was in principle entitled to general damages for a further restriction on his liberty caused by his removal to a closed category B prison.⁷⁹ The court upheld the claimant's appeal and held that the particulars of claim alleged relevant damage constituting a significant loss of liberty occasioned by his transfer from open to closed conditions. This constituted a form of material damage sufficient to found the cause of action for misfeasance in public office.⁸⁰

The facts of *Karagozlu* are similar to those of the South African case of *Zealand v Minister of Justice and Constitutional Development and Another*,⁸¹ where an awaiting trial prisoner was detained in the maximum-security wing of the prison as if he were a sentenced prisoner. The Constitutional Court unanimously

74 2007 1 WLR 1881 (CA).

75 Para 24.

76 *Roberts v Chief Constable of Cheshire Constabulary* 1999 1 WLR 662.

77 1997 3 WLR 403, 1997 2 All ER 762, 1998 QB 498 (CA).

78 1998 QB 498 (CA) 514–515.

79 See eg *R v Deputy Governor of Parkhurst Prison, Ex parte Hague* 1992 1 AC 58; *Racz v Home Office* 1994 2 AC 45 (HL).

80 2007 1 WLR 1881 (CA) 1888 paras 53 and 54.

81 2008 4 SA 458 (CC).

held that his detention was unlawful for the purpose of his claim for delictual damages in private law. There was no discussion as to whether there was financial loss or physical or material damage; the matter was simply decided on the basis of constitutional breach.⁸² In that regard, there could be no doubt as to whether injury has been caused through a breach of personal liberty; the personal deprivations arising from the wrongful confinement were sufficient to support the delict claim. Langa CJ emphasised the distinction between an awaiting trial prisoner and a convicted prisoner, something the majority of the Supreme Court of Appeal⁸³ thought was not within the ambit of what it was called upon to decide, but which Ponnann JA adverted to in his dissent.⁸⁴ The Chief Justice held that the detention of the appellant, an awaiting trial detainee, in a maximum-security facility together with sentenced and convicted prisoners, amounted to a deprivation of freedom that was arbitrary and without just cause.⁸⁵ It was in violation of s 12(1) of the 1996 Constitution.⁸⁶

8 CONCLUSION

Given the judgments in *James* and *Watkins*, it seems appropriate to ask: wither that common law sense of justice popularly expressed in the maxim: *ubi jus, ibi remedium*? In his judgment in *Ashby v Whyte*⁸⁷ from where the maxim is derived, Lord Chief Justice Holt held that: "If the plaintiff has a right, he must of necessity have means to vindicate and maintain it, and a remedy if he is injured in the exercise or enjoyment of it, and indeed it is a vain thing to imagine a right without a remedy; for want of right and want of remedy are reciprocal."⁸⁸ A careful scrutiny of the history of English law of government liability reveals a flurry of cases decided after *Ashby v Whyte* where damages were recovered against government or other public authority for unlawful exercise of public power without requiring proof of the appropriate head of tort under which the actions were brought or damage caused by the breach.⁸⁹

82 Cf *Maharaj v Attorney General of Trinidad and Tobago (2)* 1978 2 All ER 670, 1978 2 WLR 902, 1979 AC 385 (PC), where the breach was that of personal liberty.

83 *Minister of Justice and Constitutional Development v Zealand* 2007 3 All SA 588 (SCA).

84 2007 2 SA 401 (SCA), 2007 3 All SA 588 (SCA) para 28.

85 Cf per O'Regan J, *S v Coetzee* 1997 3 SA 527 (CC) para 159, speaking about the constitutional jurisprudence of the right not to be deprived of freedom arbitrarily or without just cause as having both substantive and procedural aspects. See also *De Lange v Smuts NO* 1998 3 SA 785 (CC) para 18.

86 Cf a 10(2), International Covenant on Civil and Political Rights (ICCPR). See also ss 82 and 83 of the Correctional Services Act 8 of 1959 applicable to the present case; now replaced by chapters 4 and 5, Correctional Services Act 111 of 1998 which commenced on 31 July 2004. All these laws draw significant distinctions between the two classes of prisoners in terms of privileges and restrictions necessary for the maintenance of security and good order in the prison.

87 1703 2 Raym 938.

88 953.

89 *Ibid.* See eg *Whitchurch v Hide* 1742 26 ER 636; *Ferguson v Kinnoul (Earl)* 1842 9 Cl & F 251 280; *Pickering v James* 1873 LR 8 CP 489 503; *Clifton v Hooper* 1844 ER; *Whitelegg v Richards* 1832 2 B & C 45 52; *Brayser v Maclean* 1875 LR 6 PC 398; *Dumbell v Roberts* 1944 1 All ER 326; *Beatty and Mackenzie v Kozak* 1958 13 DLR (2d) 1; *Gordon v Metropolitan Police Commissioner* 1910 2 KB 1080; *Elias v Passmore* 1934 2 KB 164; *Allen v Flood* 1898 AC 1 96; *Roncarelli v Duplessis* 1959 16 DLR (2d) 689; *McGillivray v Kimber* 1915 52 SCR 146.

Even in the absence of a written Constitution or the entrenchment of the rights adjudicated upon, the under-development of the declaratory judgment in the English public law in those early days, the fact that “the rules relating to the circumstances in which a person injured by a breach of statutory duty can recover damages form a notoriously difficult area of the law of tort”⁹⁰ and granted that actions for damages were successful in the history of public law in sporadic and limited circumstances, the courts were able to come to the conclusions they did. Now that the constitutional environment is more conducive and favourable towards judicial activism in the face of “the wider and less fettered discretion”⁹¹ expressly vested in the courts in granting constitutional remedies including damages, in modern Constitutions and Bills of Rights legislation, it follows that the courts should be more proactive than the judges of the eighteenth century.⁹²

In this regard, the approach of the courts in Ireland which has led to the development of the constitutional damages jurisprudence in that country is admirable and in the direction of the argument put forward here. The Irish experience has developed around three main propositions. First, there is the *Meskill* doctrine, which postulates that “the constitutional right carried within it, its own right to a remedy or for the enforcement of it”.⁹³ Second, “where the People by the Constitution create rights against the State or impose duties upon the State, a remedy to enforce must be deemed to be also available.”⁹⁴ In language reminiscent of that of Lord Diplock in *Maharaj*, the third proposition states that liability in these cases was not imposed on the State on vicarious liability principles but was directly and primarily imposed on the State itself acting through its constitutional organs. An action for breach of a constitutional right “is not a mere tort, it is a wrong that may defy classification in the traditional categories of wrong and is to be distinguished from instances of tortious liability for which the State may be liable vicariously for the wrongs of its employees.”⁹⁵

90 De Smith *Judicial review of administrative action* (1959) 551.

91 Per McIntyre J speaking of the intendments of s 24(1) of the *Canadian Charter of Rights and Freedoms* 1982 in *R v Mills* 1986 29 DLR (4th) 161 181. Indeed, McLachlin J said in *R v 974649 Ontario Inc* 2001 SCC 81 (CanLII), 2001 3 SCR 575, 2001 SCC 81 paras 19–20 that “a right, no matter how expansive in theory, is only as meaningful as the remedy provided for its breach”. Again, in *Doucet-Boudreau v Nova Scotia (Minister of Education)* 2003 SCC 62 (CanLII), 2003 3 SCR 3 para 25, Iacobucci and Arbour JJ held that a purposive approach to remedies provisions in the *Canadian Charter* provides, not only “a full, effective and meaningful remedy for *Charter* violations”, but also gives “modern vitality to the ancient maxim *ubi jus, ibi remedium*: where there is a right, there must be a remedy”.

92 It must be noted that in her recent exhaustive and thorough consideration of the ingredients of constitutional damages in the Canadian jurisdiction in *Vancouver (City) v Ward* 2010 SCC 27 (CanLII), MacLaughlin CJC did not mention the requirement of damage in her four-step construct. The Chief Justice of Canada stated that the following questions should be posed: (a) Has a *Charter* right been breached? (b) Are damages a just and appropriate remedy? (c) Are there countervailing factors to defeat the functional considerations that support a damage award and render damages inappropriate or unjust? (d) What quantum of damages will represent a meaningful response to the seriousness of the breach and the objectives of s 24(1) damages?

93 Per Walsh J in *Meskill v CIE* 1973 IR 121.

94 *Byrne v Ireland* 1972 IR 241 per Walsh J.

95 *Kemmy v Ireland* 2009 IEHC 178 p 6. See also *Hanrahan v Merk, Sharp and Dohme Limited* 1988 ILRM 629; *W v Ireland (No 2)* 1997 2 IR 141.

In line with these principles, courts in Ireland have established that compensation can be recovered for a breach of a constitutional right and the right to seek damages for such a breach when no other effective remedy existed. Accordingly, they have awarded damages⁹⁶ for a prisoner's right to communicate;⁹⁷ a journalist's right to privacy;⁹⁸ a right to free education;⁹⁹ right to a fair promotion's procedure;¹⁰⁰ right to liberty and good name;¹⁰¹ right to adequate primary education;¹⁰² right to fair immigration procedures;¹⁰³ and breach of electoral rights.¹⁰⁴ It follows that in the Irish courts, a breach of a constitutional right could be actionable *per se*, without proof of damage. The absence of proof of actual loss is reflected on the quantum recoverable for, in such a circumstance, the plaintiff merely recovers a nominal sum. In majority of the cases, the court would compensate as in tort action on the basis of demonstrated pecuniary and non-pecuniary loss with punitive or exemplary damages being awarded in exceptional cases.¹⁰⁵

Again, in enunciating the principles of constitutional damages in *Maharaj*, Lord Diplock emphasised the distinct nature of this remedy: it was not an action in tort. Yet, the requirement that damage must be shown to have been suffered, an element in the definition of most torts¹⁰⁶ and delict,¹⁰⁷ have been dragged into this aspect of constitutional adjudication by the Privy Council. If the separation enunciated in *Maharaj* by the same Privy Council were maintained, then, there is a significant likelihood that *James* could have been decided differently. Strikingly, the *Maharaj* jurisprudence has not been overruled and the fact that their Lordships tend not to refer to it in recent judgments, *Ramanoop*¹⁰⁸ and *Suratt*¹⁰⁹ in particular, does not lead to a conclusion that its legacy has been wiped out from the constitutional damages template of Trinidad and Tobago.

96 See also *Sinnott v Minister for Education* [2001] 2 IR 545; *O'Donoghue v Legal Aid Board* 2004 IEHC 413; *Gray v Minister for Justice* 2007 IEHC 52; *Henry v Associated Newspapers [Ireland] Ltd* 2008 IEHC 249.

97 *Kearney v Ireland* 1986 IR 116.

98 *Kennedy v Ireland* 1987 IR 587.

99 *Conway v Irish National Teacher's Association* 1991 2 IR 305.

100 *Healey v Minister for Defence*, High Court, 7 July 1994 (unreported).

101 *Walsh v Ireland* Supreme Court, 30 November 1994 (unreported).

102 *Sinnott v Minister of Education* 2001 2 IR 545.

103 *Gulyas v Minister of Justice, Equality and Law Reform* 2001 3 IR 216.

104 *Redmond v Minister for the Environment (No 2)* 2006 3 IR 1 – damages for breach by the legislative organ of state of the constitutional rights of a citizen.

105 *McFarlane v Ireland* 2010 ECHR 1272 para 85.

106 Rogers *Winfield & Jolowicz on tort* (18ed 2010) 6.2.

107 Neethling, Potgeiter and Visser, *Law of delict* (2006) 195.

108 *Attorney General of Trinidad and Tobago v Ramanoop* 2005 UKPC 15, 2005 2 WLR 1324, 2006 1 AC 328 (PC).

109 *Suratt v Attorney General of Trinidad and Tobago* 2008 UKPC 38 (PC).

Unravelling the entrapment enigma: Reflections on the role of the mental health expert in the assessment of battered woman syndrome and coercive control advanced in support of a defence of non-pathological criminal incapacity (2)*

Philip Stevens

LLB LLM

Advocate of the High Court of South Africa

Member of the Pretoria Bar

LLD candidate, University of Pretoria

4 THE PSYCHO-SOCIAL DYNAMICS OF AN ABUSIVE RELATIONSHIP: SELECTED EXAMPLES

4.1 Theory of coercive control

“It was as if he transplanted his brain into mine. I started to think like him. Dirk made me believe that we were untouchable and that we could do what we wanted⁶⁸ . . . Dirk was my god. He made me believe he was almighty and that he was in control.”⁶⁹

* See 2011 *THRHR* 432 for Part 1.

68 *Pretoria News* 11 March 2009. See the unreported judgment of Eksteen AJ in *S v Cézanne Visser* case no CC 545/09 of 27 Oct 2009 (hereafter *S v Visser*) 58.

69 *Pretoria News* 14 March 2009. These quotations were extracted from the accused’s evidence in the controversial and highly-publicised trial of *S v Visser* (*supra* fn 68) 58, often referred to as the “Barbie trial”. The facts of *Visser* were as follows: Cézanne Visser, better known as “advocate Barbie” faced fourteen charges against her which included fraud, the incitement of minors to commit indecent acts, indecent assault, rape, possession of child pornography, the manufacturing of child pornography as well as the possession of drugs. The case was heard in part in the Pretoria High Court when Patel J passed away in April 2007. The case recommenced on 16 February 2009. Visser initially stood trial on all of the charges together with her former lover and fiancé, Dirk Prinsloo. Prinsloo, however, fled the country in 2006 and was one of Interpol’s most wanted criminals. On 12 June 2009 he was arrested in Balarusia for attempted robbery of a bank. He also faced charges of theft, hooliganism and torture. On 1 February 2010 he was convicted of theft, bank robbery and hooliganism and sentenced to thirteen years in jail. Visser, who was a bright student and obtained her law degrees with distinction, moved in with Prinsloo shortly after she failed her bar examination and subsequently joined the independent bar association. Visser testified that when she met Prinsloo, she was naive and eventually got trapped in his sex-web. Some of the charges against Visser relate to occasions where she and Prinsloo performed indecent acts with minor girls they had collected from an orphanage in Pretoria for so-called “weekend-visits”. According to the evidence these girls were provided with

continued on next page

Another theory of explaining the nature and effects of domestic abuse is the theory of coercive control.⁷⁰

A domestic assault is often part of a much larger system of controlling, coercing, intimidating and violent behaviours employed by an abusive partner to control the victim. Stark indicates that according to evidence, violence in abusive relationships is an ongoing phenomenon rather than episodic in nature.⁷¹

According to Stark there is a general perception that abused women stay in an abusive relationship and develop mental health and behavioural problems due to the fact that exposure to severe forms of violence induces syndromes such as Post Traumatic Stress Disorder and Battered Woman Syndrome which prevent the woman from escaping from the abusive relationship.⁷² However, Stark indicates that only a small percentage of abuse victims display these symptoms and that many of them do not develop psychological problems.⁷³

This dominant approach to domestic violence accordingly fails to adequately address two important facts: The entrapment of victims in relationships where ongoing abuse is virtually inevitable; and the development of a problem profile that distinguishes abused women from every other class of assault victim.⁷⁴

Stark states:

“Work with battered women outside the medical complex suggests that physical violence may not be the most significant factor about most battering relationships. In all probability, the clinical profile revealed by battered women reflects the fact that they have been subjected to an ongoing strategy of intimidation, isolation and control that extends to all areas of a woman’s life, including sexuality, material necessities, relations with family, children and friends, and work. Sporadic, even severe violence makes this strategy of control effective. But the unique profile of the battered woman arises from the deprivation of liberty implied by coercion and control as it does from violence induced trauma.”⁷⁵

drinks, often “Milo”, that contained sedative drugs after which indecent sexual acts were performed on them by Prinsloo. Prinsloo also subjected Visser to various forms of brutal and sadistic sexual abuse. Visser, according to her testimony, also suffered immense psychological and emotional abuse at the hands of Prinsloo. Visser’s defence relates to the battered woman syndrome but more specifically to the coercive control that Prinsloo exerted over her. Her defence encapsulates manifestations of the Stockholm-syndrome, and the compliant victim syndrome which will be discussed below. Eksteen AJ noted (58) that Visser’s love for Prinsloo comprised of various aspects including: “A. Vrees vir die man; B. Afhanklikheid van hom; C. Die mag wat hy oor haar gehad het” and noted (63) the following in respect of Visser’s defence: “Sy beweer sy was vasgevang in Dirk Prinsloo se mag. Sy het in haar binneste gevoel dit is nie reg nie, maar dit was nie sterk genoeg om teen Prinsloo se mag in opstand te kom nie. ‘Ek het nie gedink wanneer hierdie goed gebeur nie’ het beskuldigde gesê.”

70 Stark (2007) 112; Ludsin and Vetten 67.

71 Herman (1992) 74; Stark “Symposium on reconceptualizing violence against women by intimate partners” 1995 *Albany LR* 973; Stark “Coercive control” (justiceformothers.com/documents/coercivecontrol.pdf; accessed 23-03-2009). See also Okun *Woman abuse: Facts replacing myths* (1986) 86.

72 *Ibid.*

73 *Ibid.*

74 *Ibid.*

75 1995 *Albany LR* as quoted by Ludsin and Vetten 61–62. See also Stark (2007) 57 where he states that protecting women from severe assaults has resulted in the substitution of physical abuse with coercive control.

Coercive control is a model of abuse that includes various strategies employed by an abusive partner to dominate women in their personal life.⁷⁶ According to Stark, coercive control “describes an ongoing pattern of sexual mastery by which abusive partners, almost exclusively males, interweave physical abuse with three equally important tactics: intimation, isolation and control”.⁷⁷

Stark notes that it is important to distinguish the coercive control model from the traditional domestic violence model and highlights the following essential differences:

- The domestic violence model typically emphasizes the familial, cultural and psychological foundations of abusive behaviour. The coercive control model, on the other hand, views the dynamics in abusive relations against the backdrop of the historical battle for women’s liberation and men’s motivation for preserving their traditional privileges in personal life amidst this battle.
- Domestic violence laws generally follow an incident-specific approach and accordingly measure the severity of abuse against the level of force or injuries inflicted. The coercive control model, on the other hand, is predicated on the premise that most battered women who seek help experience “coercion” as “ongoing” rather than merely “repeated”, and that the most important aspect of these assaults lies in their frequency or even their “routine” nature, rather than their severity. This results in a “cumulative” effect found in no other assault crime.
- Physical harm and psychological trauma that are generally very important phenomena in the domestic violence model, remain important in the coercive control model, but its theory of harms substitutes the violation of physical integrity with an emphasis on infringements of “liberty” that includes the deprivation of rights and resources essential to personal autonomy.⁷⁸
- In the coercive control model, what men do to women is less important than what they prevent women from doing for themselves.

With increasing efforts by women to ensure their equality in a previously-male-dominant society, men find it more difficult to ensure women’s obedience and dependence through the application of violence alone. Accordingly, in the face of reality men have expanded their oppressive techniques to encompass a range of constraints on women’s autonomy formerly imposed by law, religion, and women’s exclusion from the economic, cultural and political mainstream, “in essence trying to construct a ‘patriarchy in nature’ in each individual relationship, the course of malevolent conduct known as coercive control”.⁷⁹

The theory of coercive control originated in the experiences of people who lived in situations of captivity or people who were taken hostage and displayed symptoms of the “Stockholm-syndrome” or “traumatic bonding”.⁸⁰

76 Stark “Coercive control” (fn 71 *supra*). Coercive control is also sometimes referred to as coerced persuasion; conjugal, patriarchal or intimate terrorism, emotional or psychological abuse; indirect abuse or emotional torture.

77 *Ibid.* See also Stark (2007) 5.

78 *Ibid.*

79 *Ibid.*

80 Ludsin and Vetten 67; Stark (2007) 198 203. The concepts of “Stockholm-syndrome” and “traumatic bonding” are discussed below.

According to Herman, captivity “brings the victim into prolonged contact with the perpetrator, and creates a special type of relationship, one of coercive control”.⁸¹ The motivations for this behaviour are the following:⁸² Complete control over the victim; making the victim acquiesce in her domination; and due to the fact that it appears that the victim accepts this abusive treatment, the likelihood of outside assistance is reduced.

Coercive control can also in some cases result in the victim identifying with the abuser. The latter entails that the abused woman will attempt to view the world through the eyes of the abuser in an attempt to prevent further harm and danger.⁸³

Stark states⁸⁴ that the combination of these big and little indignities most adequately explains why women suffer and respond as they do within abusive relationships, and also why so many women become entrapped, why some battered women kill their partners as well as the reasons as to why they are prone to develop a range of psychosocial problems and exhibit behaviours that are contrary to basic common sense.

4.1.1 Methodology and techniques of coercive control

Ludsin and Vetten state that through the systematic and repetitive infliction of psychological trauma together with violence, terror and helplessness become part of the abused woman and her sense of self is slowly eliminated.⁸⁵

According to Stark, coercion implies the use of force or threats to compel or dispel a particular response.⁸⁶ Accordingly, apart from inflicting immediate pain, injury, fear or death, coercion also causes long-term physical, behavioural or psychological trauma. Control consists of various forms of deprivation, exploitation and command that compel obedience in an indirect manner by monopolising important resources, dictating choices, micro-regulating the woman’s behaviour, restricting her options and depriving her of support essential to the execution of independent judgment.⁸⁷

Stark notes⁸⁸ that control may be implemented by means of specific acts of prohibition such as keeping the victim from going to work or denying the victim access to a car or phone.

81 Ludsin and Vetten 67; Herman 74.

82 *Ibid.*

83 Ludsin and Vetten 67. They note that the essential difference between abused women and prisoners is that women do not need to be captured or detained within the relationship as they remain within the relationship out of love for the abusive partner. Accordingly they will be less inclined to offer resistance than other captives as a result of their commitment to the relationship. The shocking reality, however, is that when affection has faded, many women may remain captive through economic dependence, physical force and social, psychological and legal subordination. See also Herman 74.

84 (2007) 15–16.

85 Ludsin and Vetten 68. See also Okun 119.

86 Stark (2007) 228.

87 *Ibid.*

88 *Ibid.* Stark in addition notes (228): “Control may be implemented through specific acts of prohibition or coercion, as when a victim is kept home from work, denied access to a car or phone, or forced to turn over her paycheck. But its link to dependence and/or obedience is usually more distal than coercion and so harder to detect, making assigning responsibility a matter of working back from its effects through a complex chain of prior events. The result

continued on next page

Despite the fact that violence is one of the essential features of coercive control, it need not necessarily recur constantly or in the same brutal degree on every occasion.⁸⁹ An abusive partner can use violence only when necessary and to the extent that is required to instil fear and obedience in the victim. Thereafter mere threats of violence will suffice to render the woman compliant.⁹⁰

Threats of violence may also sometimes extend to the abused woman's children and her family and result in fear and anxiety.⁹¹ Ludsin and Vetten state that, in order to understand what will make an abusive partner happy and so reduce further harm and violence, the abused woman often attempts to get inside the abuser's head.⁹²

Accordingly the woman not only becomes sensitive to the abuser's moods and whims but submissive to his demands. In attempting to view aspects in the light that the abuser does, the abused woman adopts the abuser's outlook and believes that she is the cause of the abuse and that she deserves it.⁹³ This is also referred to as the "identification with the aggressor" and eventually results in the woman's identity being defined by the abusive partner.⁹⁴

Severe emotional and psychological abuse breaks down the woman's personality and eventually makes her believe that such degradation defines who she is. According to Ludsin and Vetten, women are often induced into taking responsibility for the abuse by means of forcing the woman to make false "confessions".⁹⁵ These include forcing the woman to admit to transgressions, non-existent sexual relationships or even a confession that the woman is to be blamed for the abuse. These admissions are sought to justify the abuser's abuse and results in further breakdown in the woman's identity and self-esteem, causing her to feel responsible and deserving of abuse.⁹⁶

when coercion and control are combined is the condition of unreciprocated authority often identified as domination and victims experience as entrapment."

89 Ludsin and Vetten 68.

90 *Ibid*; Herman 77. See also Stark (2007) 250 who states that threats violate the person's right to physical and psychic security. Stark notes that in an English study, 79,5% of the women reported that their partner threatened to kill them at least once; and 43,8% did so often. In addition 60% of the men threatened to have the children taken away; 36% threatened to hurt the children; 32 % threatened to have the woman committed; 82% threatened to destroy things that the woman cared for.

91 Ludsin and Vetten 69.

92 *Ibid*.

93 *Ibid*.

94 *Ibid*. See also Graham, Rawlings and Rigsby *Loving to survive: Sexual terror, men's violence and women's lives* (1994) 37–39.

95 *Ibid*.

96 Ludsin and Vetten 70. According to them, these forced confessions are often accompanied by other methods employed to instil a misplaced guilt in the victim with the aim of rendering her to feel responsible for the abuse, including: Guilt by association in the sense that the woman is scorned if she associates with people the abuser does not approve of, such as friends or family; guilt by intention in holding the woman culpable for having motives which could result in behaviour harmful to the abuser; guilt for negative attitudes towards the abuser or for doubting his decisions; guilt for having knowledge which could incriminate the abuser; guilt for taking actions which is harmful to the abuser, regardless of the fact that harm was not intended, for example where the woman was delayed on her way home and could not prepare dinner; guilt for not supporting the coercive controller's interests – in these relationships women are expected to stand up and abide their abusive partners, regardless of the nature of the abuser's actions; and guilt for personal faults. See also Stark (2007) 250–271; Okun 130–132.

In terms of coercive control, sexual coercion is often a prominent feature of the exercise of control. Stark notes that women are often forced to have sex against their will often or all the time.⁹⁷ According to Ludsin and Vetten sexual abuse serves not only as a means of control, but also constitutes a form of degradation.⁹⁸ In situations of coercive control, an abuser may also regulate what the woman eats, when she sleeps and what she should wear.⁹⁹ Coercive control can also comprise of various monitoring and surveillance tactics which may include phoning or arriving at the woman's place of work in order to ensure that she is indeed at work or checking the calls made from her cell phone.¹⁰⁰

Isolation is another prominent feature of the coercive control model. Controllers generally isolate their partners with the aim of preventing disclosure, instilling dependence and also to restrict the woman's skills and resources in order to prevent her from seeking help and support.¹⁰¹ The abusive partner will accordingly isolate the woman from her friends, family and other support systems and limit her contact with others to those who support the controller's or abuser's perspectives.¹⁰²

97 Stark (2007) 243. See also *S v Visser supra* fn 69 where Visser testified that her ex-lover, Prinsloo, had sex on various occasions per day. See *Beeld* and *Pretoria News* 11 March 2009. Visser stated: "Ek weet nie hoe hy dit kon regkry nie, maar dit was baie."

98 Ludsin and Vetten 71; Herman 79.

99 Ludsin and Vetten 72. See also *S v Visser supra* fn 69 in *Beeld* and *Pretoria News* 27 February 2009 where Visser testified that when she and her former lover, Prinsloo, went out, he laid out the skimpy outfits she had to wear. She also testified that while he had wholesome meals, she was forced to go to the gym and lived on protein shakes as Prinsloo hated cellulite.

100 Ludsin and Vetten 72; Stark (2007) 255. Stark notes that in coercive control surveillance entails a range of monitoring tactics and is aimed at establishing that the abuser is omnipotent and omnipresent.

101 Stark (2007) 262. See also *S v Ferreira (supra* fn 6) paras 29 and 35.

102 Ludsin and Vetten 72. See also *S v Engelbrecht supra* fn 6 where the accused was charged with murder of her abusive husband. The evidence revealed that the deceased stalked the accused, physically assaulted and psychologically and emotionally abused her on a regular basis. The deceased also repeatedly threatened to kill her. The accused eventually thumb-cuffed the deceased while he was sleeping and suffocated him by placing a plastic bag over his head which she tied with the belt of her dressing gown. He died as a result of suffocation. The defence relied upon by the accused was not non-pathological criminal incapacity but expert evidence pertaining to the battered woman syndrome was presented. The two experts who presented expert evidence in this case were Mr Leonard Carr, a clinical psychologist and Ms Lisa Vetten, Gender Coordinator for the Centre for the Study of Violence and Reconciliation. Ms Vetten also explained the theory of coercive control and testified that the accused felt like a prisoner in her own home and that the deceased's constant invasion in her life and work resulted in a "monopolisation of her perceptions" (paras 181 and 182). Ms Vetten testified: "So successful were the deceased's attempts to control Mrs Engelbrecht that she began restricting her own activities. She avoided speaking to men, including her neighbours, for fear the deceased would suspect something was going on . . . The deceased did not need to assault Mrs Engelbrecht very severely or very often. Those occasions when he did use violence were sufficient to instil fear in Mrs Engelbrecht" (paras 185 and 186). Ms Vetten also testified (para 200) that the deceased abused the accused physically, sexually, verbally and psychologically and the abuse intensified during the course of the relationship. According to Ms Vetten, the pattern of coercion and control to which the accused was subjected, extended to every aspect of their existence, resulting in her isolation and entrapment within the relationship. The accused had been depersonalised and dehumanised, experiencing herself as no more than a thing (para 201).

Eventually the abusive partner becomes omnipotent and omnipresent with complete power over the abused woman. Herman encapsulates this control in the following dramatic terms:¹⁰³

“The repeated experience of terror and reprieve, especially within the isolated context of a love relationship, may result in a feeling of intense, almost worshipful dependence upon an all powerful, godlike authority. The victim may live in terror of his wrath, but she may also view him as the source of strength, guidance and life itself. The relationship may take on an extraordinary quality of specialness. Some battered women speak of entering a kind of exclusive, almost delusional world, embracing the grandiose belief system of their mates and voluntarily suppressing their own doubts as proof of loyalty and submission.”

From the above discussion it is clear that expert evidence becomes a crucial tool in comprehending the battered woman’s actions. The coercive control theory of abuse constitutes one of those complex areas where expert evidence becomes inescapable.

4.2 The Stockholm syndrome

Another theory used to explain powerful emotional attachments between an abused woman and her abusive partner is the so-called Stockholm syndrome.¹⁰⁴

The Stockholm syndrome is a psychological response often observed in abducted hostages, in which the hostage displays signs of loyalty to the hostage taker, regardless of the danger or risk in which they find themselves. The syndrome is named after the Norrmalmstorg robbery of the *kreditbanken* at Norrmalmstorg, Stockholm, Sweden, in which the bank robbers held bank employees hostage from 23 to 28 August 1973. The victims became emotionally attached to their victimisers and eventually even defended their captors after they were freed from their six-day ordeal.¹⁰⁵

In terms of the Stockholm syndrome, captives begin to identify with their captors initially as a defensive mechanism, out of fear of further violence. Stockholm syndrome is also commonly encountered in abusive relationships and is accordingly used as a model for explaining why an abused woman did not leave her abusive husband.

Every syndrome has symptoms or behaviours, and Stockholm syndrome is no exception. While a clear-cut list has not been established due to varying opinions by researchers and experts, several of these features will be present:¹⁰⁶ Positive feelings by the victim toward the abuser/controller; negative feelings by the victim toward family, friends or authorities trying to rescue/support them or win their release; support for the abuser’s reasons and behaviours; positive feelings by the abuser toward the victim; supportive behaviours by the victim, at times helping the abuser; and the inability to engage in behaviours that may assist in their release or detachment.

103 Herman as quoted in Ludsin and Vetten 72–73.

104 Wikipedia encyclopedia <http://en-wikipedia.org/wiki/stockholm-syndrome> (accessed 2009-02-23); “Love and Stockholm syndrome: The mystery of loving an abuser” <http://counsellingresource.com/quizzes/stockholm/index/html> (accessed 2009-02-23); Ludsin and Vetten 73–74.

105 *Ibid.* The term “Stockholm syndrome” was coined by the criminologist and psychiatrist Nils Bejerot, who provided assistance to the police during the robbery.

106 “Love and Stockholm Syndrome” *supra* fn 104.

It has been found that four situations or conditions are present to serve as a foundation for the development of Stockholm syndrome. These four situations can be found in hostage, severe abuse and abusive relationships:¹⁰⁷ The presence of a perceived threat to one's physical or psychological survival and the belief that the abuser would carry out the threat; the presence of a perceived small kindness from the abuser to the victim; isolation from perspectives other than those of the abuser; and the perceived inability to escape the situation.

According to Ludsin and Vetten this traumatic bond develops over time and by the time women realise that the abuse is inescapable, the emotional bond created by the domestic violence is very strong.¹⁰⁸ The Stockholm syndrome produces an unhealthy bond with the controller and abuser. It is the reason many victims continue to support an abuser even after a relationship has ended. It could also be used to explain why abused women continue to see "the good side" of an abuser and appear sympathetic to someone who has mentally and most often psychologically abused them.

4.3 Compliant victim of the sexual sadist

Abused women involved in sexual and violent crimes have been referred to as "compliant victims" in order to illustrate their submissive cooperation in their own and others' victimization.¹⁰⁹ These relationships are typified by the most brutal forms of sexual violence and comprise a complete transformation of the woman's sense of self, also in respect of her behaviour in response to intimate contact, sexual fantasies and the desires of the sadistic male.¹¹⁰ Hazelwood, Warren and Dietz state that the battered women of a sexual sadists experience a process of coercion similar to brainwashing.¹¹¹ Accordingly these women experience a process of manipulation of various rewards and punishments within a context of social isolation which "can alter self concept, expectations, and behaviours among at least some victims".¹¹² To the average lay observer these women seem to be experiencing abuse despite the opportunities they have of escaping.¹¹³

Hazelwood, Warren and Dietz state that in this form of domestic violence the "captor" seeks compliance as well as opportunities for continued abuse.¹¹⁴ Chapman indicates that the motivation for the woman to submit herself to the acts of the abuser is not simply to please him and in some instances the women become assimilated into the sexual aggression of their partners.¹¹⁵ They further

107 *Ibid.*

108 Ludsin and Vetten 74; Dutton and Painter "Emotional attachments in abusive relationships: A test of traumatic bonding theory" 1993 *Violence and Victims* 105.

109 Chapman "The compliant victim of the sexual sadist and the proposed Canadian defence of coercive persuasion" (paper submitted at the International Society for the Reform of Criminal Law – 22nd International Conference Codifying the Criminal Law: Modern Initiatives (2008) Dublin, Ireland) 15. See also Warren and Hazelwood "Relational patterns associated with sexual sadism: A study of 20 wives and girlfriends" 2002 *J of Family Violence* 75–77.

110 *Ibid.*

111 "Compliant victims of the sexual sadist" 1993 *Australian Family Physician* 474.

112 *Ibid.* See also Chapman (2008) 15.

113 *Ibid.*

114 *Ibid.*

115 Chapman (2008) 16.

note that the woman's response to the paraphilic interest of the man could be conceptualised by the gradual assimilation of behaviour that integrates the sadist's sexual fantasies into her own behaviour.¹¹⁶

What is striking is that most women within these abusive relationships are successful professionally when they meet the abuser.¹¹⁷ Sexual sadists prefer professional women as they have the desire to prove that they can transform a woman from a nice middle class family and reduce her to a "sexual slave", willing to join them in any act no matter how degrading or humiliating.¹¹⁸ These relationships are also categorized with physical, emotional, psychological and sexual abuse.¹¹⁹

A further intrinsic and prominent feature of abuse within this context relates to the process of transformation that women undergo from relatively normal patterns of living to complete bizarre, destructive and dangerous forms of exploitation and perversion.¹²⁰

There is a striking pattern of coercive persuasion among women within these relationships of abuse. Five factors are present in most women:¹²¹

- *Selection of vulnerable women* – the men generally sought naive, passive and vulnerable women whom the sadist would be able to use for his own need for dominance, control and sexual desires.
- *Seduction of the targeted woman* – Hazelwood, Warren and Dietz state that all of the women they studied indicated that the abuser was charming, considerate and unselfish when they met and all of the women entered into the relationship quickly regardless of the fact that they identified a sinister side to the abusers.
- *Shaping sexual behaviour* – shaping of the woman's sexual behaviour was dependent on readiness of the woman to engage in alternative sexual acts, and the abuser typically expresses gratitude for the participation in these activities or disappointment if she did not participate. The sexual sadist typically persuades the woman to engage in sexual activity beyond her normal repertoire.
- *Social isolation* – the sadist becomes possessive and jealous of activities that do not include him and rejects the woman's friends and family, thereby isolating her.
- *Punishment* – psychological and physical abuse constitutes the final step in the transformation process.

116 474.

117 *Idem* 99.

118 Chapman (2008) 20.

119 *Idem* 20–21 notes that the research suggests that the physical abuse of women in these relationships is shocking. A study conducted on various women in these relationships showed that most of them were frequently beaten by the abuser or with objects. One woman was tied with adhesive tape over her entire body while being beaten. All of the women were sexually abused, sometimes with foreign objects. A wide range of degrading acts were performed on them, including ejaculating on their face or mouth, being urinated on, forced enemas, sex with third parties and sex with kidnapped parties. All of the women suffered psychological as well as emotional abuse. All of the women were verbally abused in order to lower their self-esteem.

120 Hazelwood, Warren and Dietz 99.

121 Chapman (2008) 21–22; Hazelwood, Warren and Dietz 99–100; Hazelwood and Michaud *Dark dreams* (2001) 112–113.

The degradation, humiliation and emotional and psychological suffering that women within these abusive situations endure, “illustrate . . . the exploitative and inhumane behaviour that one person can intentionally inflict on another”.¹²²

The problematic issue pertaining to the compliant victim is that she remains compliant for so long despite the abuse. In most cases of the compliant victim, the sadist selected a woman of higher status and transformed her into a sexually and psychologically compliant slave.¹²³ Vulnerable women are prime targets for the sexual sadist. The sadist isolates the woman, physically abuses her, deprives her of sleep, degrades and humiliates her.¹²⁴ Hazelwood, Warren and Dietz similarly conclude: “The pleasure in complete domination over another person is the very essence of the sadistic drive.”

Expert evidence in cases of this nature will be crucial to explain why the woman engaged in criminal activities with the abusive partner who can also very well be a sexual sadist. The coercive control model for explaining attachment and bonding within an abusive relationship has always been present, yet it is relatively new in terms of the recognition of this theory in explaining the behaviour of abused women.

The theory of coercive control in the face of the compliant victim of the sexual sadist is also currently a highly controversial aspect within the South African context.

When applying the abovementioned principles pertaining to the abused woman as a compliant victim of sexual sadistic abuse, there are striking resemblances to the evidence tendered in the highly controversial trial of Cézanne Visser.¹²⁵ The essential facts of this case are as follows:

Cézanne Visser (“Visser”) met Dirk Prinsloo (“Prinsloo”) at the age of 23. Before meeting Prinsloo, Visser failed her bar examination and was also the victim of a failed relationship. She moved in with Prinsloo soon after meeting him and a romantic relationship arose between them. Initially Prinsloo showered Visser with gifts and compliments and also referred to her as his “princess”. Soon after Visser moved in with Prinsloo she was exposed to having oral sex with him. At that stage she was still naive as to aspects involving sex, but thought that what Prinsloo expected of her was normal. At the time when she met Prinsloo, Visser testified that she had an extremely low self-esteem and was very vulnerable. Prinsloo soon displayed signs of sex addiction. Visser testified that Prinsloo watched pornographic films every morning at breakfast. Prinsloo requested Visser to have tattoo’s engraved on her body and also to have breast enlargements. Prinsloo did not get on well with Visser’s parents and eventually, on his demand, they got a protection order in terms of the Domestic Violence Act¹²⁶ against Visser’s parents which prohibited any means of contact between her and her parents. Prinsloo also controlled what Visser ate and the clothes she wore. Prinsloo requested that Visser go on a protein-shake diet to the exclusion of all other food, as he “hated cellulite”. When asked why she tolerated this, Visser testified “I wanted to please Dirk”. Visser testified that sex with Prinsloo

122 Hazelwood, Warren and Dietz 100.

123 *Ibid.*

124 *Ibid.*

125 *S v Visser supra* fn 69.

126 116 of 1998.

was brutal and without love and that she was subjected to degrading acts with foreign objects such as cucumbers, carrots and a firearm and was also forced into acts with dogs. Visser testified that she was once forced to have sex in a chapel. Visser stated:¹²⁷ “Hy het bo-op my geklim en seks met my gehad. Dit was brutaal. Hy het my hare gepluk terwyl hy my in die gesig gespoeg het.”

Visser testified that Prinsloo instructed her to have threesome sessions with other women and that she (Visser) had to recruit prostitutes for him (Prinsloo). Prinsloo also provided Visser with a book called *The story of O* which was about a man who subjected his wife to bizarre and brutal sex. She (Visser) also had to pierce her body and she testified:¹²⁸ “Dirk thought it was pretty. It was how his slut had to look.”

Prinsloo fantasised about having sex with young girls. Some of the charges against Prinsloo and Visser related to young girls they fetched from an orphanage under the guise of wanting to treat the girls for weekend-visits at their home. One girl was 15 years old; the other girl 11. The charges relate to various sexual offences including indecent assault, rape and incitement of a minor into sexual acts.

Visser admitted to performing oral sex on Prinsloo in front of the 15 year old girl and having sex with Prinsloo in front of the 11 year old girl. When asked why she performed these acts, Visser stated:¹²⁹ “I said it before and I say it now, Dirk spoke and I did. I have no idea why, but that is how it was.”

Prinsloo subjected Visser to various forms of degradation such as to drink his urine and smear his faeces on her. The latter was her punishment if she did not behave as Prinsloo wanted her to. Visser also testified that Prinsloo was addicted to sex. Accordingly, Prinsloo displayed signs of sexual sadism, whilst Visser displayed signs of the typical compliant victim. Visser also testified that kinky sexual acts were as normal as “brushing teeth”. Visser testified that she was a victim in a relationship which was not normal.

In support of Visser’s defence that she suffered from Battered Woman Syndrome, Professor Jonathan Scholtz, head of clinical psychology at Weskoppies Hospital, testified that although Prinsloo initially appeared to be Visser’s knight in shining armour, charming her and purporting to save her from an abusive family situation, he systematically and deliberately took control of her and shaped her to his needs. Scholtz further stated that Visser’s parents’ unhappy marriage, the values instilled in her during childhood and her low self-esteem made her the perfect target for Prinsloo.

Scholtz testified that while Visser was highly intelligent, she was naive and mentally immature. According to Scholtz, Prinsloo took complete control over Visser – her appearance, what she ate, when she slept and with whom she spoke. She was exposed to perverse acts with multiple people. Scholtz testified that in his opinion, Visser was subjected to severe domestic abuse and coercive control. Scholtz indicated that Prinsloo was a sexual sadist, a paedophile and suffered from other sexual deviations. According to Scholtz sexual sadists often displayed

127 *Beeld* 27 February 2009.

128 *Ibid*; *Pretoria News* 27 February 2009.

129 *Pretoria News* 10 March 2009.

various disorders, collected pornography and had serious personality disorders such as narcissism.¹³⁰

Visser was subsequently convicted on eleven of the fourteen charges on 6 and 7 October 2009.¹³¹ In delivering judgment, Eksteen AJ ultimately rejected Visser's defence of having been under the "spell" or coercive control of Prinsloo.¹³² It was also held that Prinsloo was often not present when Visser exposed herself to the children.¹³³ Eksteen AJ found that Visser did not follow everything Prinsloo instructed her to do and held that Visser sought to hide behind Prinsloo's conduct to justify her own conduct with specific reference to Visser claiming that Prinsloo was manipulative and as a result she had no will of her own.¹³⁴

130 *Beeld* and *Pretoria News* 27 March 2009, 28 March 2009; *Rapport* 5 April 2009. The author sourced the information from his personal attendance of the trial. See 79 of the unreported judgment of *S v Visser supra* fn 69 where it is noted that Prof Scholtz found that Visser was exposed to "severe domestic abuse and coercive control" and her capacity to act in accordance with her appreciation of the wrongfulness of her actions was severely compromised. Prof Scholtz in addition elaborated on the phenomenon of the battered woman syndrome and noted that it comprises more than mere physical violence and also included coercive control and intimidation. Prof Scholtz described the bond which develops between the woman and her aggressor as "ambivalent" and noted that such bond comprises two components: "A. Die vrees vir die magsoeker; B. Die adoratie of verliefdheid". Prof Scholtz described this bond as "traumatic bonding" or "paradoxical attachment" and defined coercive control (81) as "die proses waartydens die wil van die vrou onderwerp word aan die wil van die man. 'n Vrou in die situasie van vrees of dwang en belonging kan iets doen sonder dat die teenwoordigheid van die magsoeker 'n vereiste is om 'n handeling uit te voer".

131 Visser was convicted on counts of fraud; three counts of soliciting a fifteen year old to commit indecent acts by showing the child her private parts, by showing her pornography and conducting sexually-explicit conversations with her; one count of indecent assault of an eleven year-old orphan by showing pornography to her, by demonstrating to her how a vibrator worked, exposing herself to the child and having sex with Prinsloo in her presence; one count of being a beneficiary to the indecent assault of a twenty year old woman; one count of indecent assault of an adult woman by fondling her breasts and private parts and suggesting that she have sex with her and Prinsloo; one count of indecent assault on a fourteen year old drug addict who was drugged and fondled; one count of indecent assault of a twenty year old woman who was fondled after being given drugs which induced drowsiness; one count of possession of child pornography; one count of manufacturing child pornography relating to her committing indecent acts on a child, of which pictures were taken. Visser was acquitted on one count of indecent assault on an eleven year old as it was held that the child was not forced to take off her clothes during an incident at the swimming pool of the residence of Prinsloo and Visser; one count of possession of 13.2g of dagga and one count of manufacturing child pornography pertaining to a fourteen year old girl as it could not be ascertained whether pictures were indeed taken. See also *Pretoria News* and *Beeld* of 8 October 2009. It is to be noted that at the stage of completion of this article, the *Visser* judgment had not yet been reported.

132 *Ibid.* See *S v Visser (supra* fn 69) 110 Eksteen AJ held: "Daar is geen sprake dat beskuldigde willoos gehandel het nie. Die hof het geen twyfel om vanweë die inherente onwaarskynlikhede en onbetroubaarheid van haar relaas, beskuldigde se weergawe as vals te verwerp."

133 *Ibid.*

134 *Ibid.* Eksteen AJ stated: "She accepted little or any of the blame herself. She is, everytime, the victim of the conduct of others. She blames everyone else, except herself" (*Pretoria News* 8 October 2009). See *S v Visser (supra* fn 69) 95 where Eksteen AJ stated that "min of enige blaam aanvaar beskuldigde self. Sy is telkens die slagoffer van ander se optrede".

Eksteen AJ held that it was improbable that Visser was caught in Prinsloo's web as she had freedom of movement and there was also evidence to the effect that Visser could stand up to Prinsloo.¹³⁵ Eksteen AJ held as follows:¹³⁶

- Visser was a willing partner in the sexual abuse of the three children and the three young women;
- Visser willingly participated in the various sexual acts perpetrated on the victims at the Prinsloo home and she embraced the new life Prinsloo offered her;
- Visser took the initiative in locating some of the victims;
- Visser's conduct was aimed at obtaining children and women to sexually abuse for her and Prinsloo's own gain;
- Visser and Prinsloo had sex in front of some of the children to solicit them to commit indecent acts;
- Visser was aware of the fact that Prinsloo used medication to drug some of the victims and that their drinks were spiked;
- many of the acts against the children were committed in the absence of Prinsloo;¹³⁷
- Visser went to an orphanage and informed management that she and Prinsloo were married in order to persuade them to allow the children to visit them for weekends;
- Visser took the children home well-knowing what fate awaited them there.

On 24 February 2010, Visser was sentenced to an effective term of seven years imprisonment.¹³⁸ Leave to appeal to the Supreme Court of Appeal in Bloemfontein against her conviction was rejected on 13 May 2010. Visser currently resumed serving her sentence.¹³⁹

Visser not only gave rise to immense publicity, but also shed light on the age-old phenomenon of abuse within an intimate relationship. The aspects of coercive control, Stockholm syndrome and the compliant victim of the sexual sadist were brought to the fore. Even though their value within this decision remains

135 *Ibid.* Eksteen AJ specifically referred to the evidence of one of the witness, Laurie Pieters, a former friend of Visser, who on various occasions offered help to Visser, specifically with reference to Pieters offering Visser a home to stay. Pieters in addition testified that Prinsloo "was a coward and a bully. His bark was worse than his bite". See 92–94 of the unreported judgment. See also 97–100 where Eksteen AJ states that Visser's freedom of movement was not constricted by Prinsloo and that it was improbable that she was caught in his "web".

136 *Ibid.* See 124–139.

137 *Ibid.* These acts include the following: Demonstrating the use of a vibrator on herself in front of an eleven year old; fondling the child's private parts; inviting the child to remove all her clothes at the swimming pool; offering to have the child's private parts waxed; Visser also insisted on her own initiative to take the eleven year old home for a weekend visit; and Visser also instructed the children not to tell anyone about what happened to them. See also *S v Visser (supra* fn 69) 61.

138 See *Pretoria News* and *Beeld* 25 February 2010; and *Rapport* 28 February 2010.

139 See *Pretoria News* 14 May 2010 and *Pretoria News* 17 May 2010. For purposes of this article, expert evidence adduced during the sentencing of Visser is not addressed. Eksteen AJ refused to grant leave to appeal on the merits and Visser's legal team had to petition to the Supreme Court of Appeal which petition failed.

dubious, the emphasis placed on these phenomena could be seen as a positive step towards taking cognisance not only of the visible or physical aspects of abuse within intimate relationships, but also the invisible and often concealed forms abuse conducted behind closed doors. Abuse encapsulates numerous manifestations of which physical abuse is but one example. It is crucial to also acknowledge the various other manifestations of abuse such as coercive control, the Stockholm syndrome and the compliant victim theory as these are manifestations frequently encountered within abusive relationships often underscored for its impact and intensity.¹⁴⁰ Every abusive relationship will have its own distinctive characteristics distinguishing it from other abusive relationships.

The fact that the defence as put forward in *Visser* was rejected should not be construed as closing the door to this defence in cases where abused partners subjected to coercive control and manipulation commit crimes whilst under such control or manipulation. What becomes abundantly clear is that courts will approach such defences with circumspection, thus necessitating the need for effective expert testimony in support thereof. It could further be argued that the defence in *Visser* contributed to mitigation of sentence.

The distinguishing factor of *Visser* as opposed to other cases dealing with abuse, such as *Engelbrecht* and *Ferreira*, is the fact that Visser's actions were not directed against her abuser, but primarily against innocent third parties or victims. Usually within an abusive relationship the abused partner retreats and directs his or her actions against the abuser and often kills the latter. The conduct of Prinsloo towards Visser was, however, at certain stages so vile and shocking, that it could be argued that the probabilities indicate that she must have been subjected to some form of control by Prinsloo.¹⁴¹ Whether such control was of such a nature and degree as to render Visser powerless to Prinsloo, however, remains questionable. The undeniable fact is, however, that no matter how much empathy a court retains for a victim of abuse such as Visser, the court also has a duty of upholding justice for the innocent victims and of protecting the interests of minor children and their right to be protected from all forms of sexual abuse. It is submitted that the latter could be construed as one of the overriding factors negating Visser's defence.

That is why Visser and Prinsloo fit the profile of the sexual sadist and compliant victim discussed above. Whether a defence founded on these principles will succeed will depend on the circumstances and evidence presented.

140 See also *S v Engelbrecht supra* fn 6 and *S v Ferreira supra* fn 6 where traits of coercive control were present.

141 Visser's testimony to this effect included that Prinsloo on various occasions forced her to take part in threesomes with prostitutes; Prinsloo inserted various objects into her private parts; Prinsloo forced Visser to have sex in a chapel and swear at God whilst spitting in her face; Prinsloo forced Visser to perform acts on dogs; Prinsloo did not hesitate to ejaculate in her mouth; Prinsloo urinated in her mouth and forced her, in order to obtain forgiveness, to collect his faeces and rub it over her body and lick it off her hands. These are examples of some of the horrific acts Visser was subjected to. It could be argued that a woman in her sound and sober senses would not allow such acts to be performed on her. In *S v Visser (supra* fn 69) 121 Eksteen AJ also noted: "Die aanvaarde getuienis binne en buite die beskuldigde se relaas dui onomwonde dat Prinsloo moontlik met 'n proses van isolasie of 'coercive control' besig was. Die aanvaarde getuienis dui na die Hof se oordeel nie daarop dat beskuldigde tydens die pleeg van die ten laste-gelegde misdrywe ten volle onder die mag of dominansie van Prinsloo was nie."

5 ROLE OF EXPERT EVIDENCE IN CASES OF BATTERED WOMAN SYNDROME

Central to all the theories explaining why an abused woman reacted to abuse in the particular way she did, stands the mental health expert called to assess the battered woman. Within the context of the battered woman who eventually kills her abusive partner, the mental health expert who will typically be a psychologist or a psychiatrist will have to assist the court in explaining the battered woman's dilemma and why she eventually resorted to deadly force instead of exploring alternative options. Expert evidence is generally presented to combat existing myths about battered women and not to address the ultimate issue of guilt or innocence. Expert evidence on the battered woman syndrome entails the psychological traits that typify battered women as well as their perceptions of the potential dangerousness of the abuser's possible violence.¹⁴²

Hudsmith¹⁴³ notes that as a result of the non-traditional nature of a battered woman's resort to the use of deadly force, the reasonableness of her perceptions of danger may not always be transparent. Expert evidence is accordingly crucial to explain the dynamics of the abusive relationship and the effect the violence may have on a battered woman's perceptions of danger.¹⁴⁴

142 Blackman "Potential uses for expert testimony: Ideas toward the representation of battered women who kill" 1986 *Women's Rights Law Reporter* 227-228. See also Walker (1979) 56-70; Ludsin and Vetten 93; Reddi 261; Johann and Osanka (1989) 157; Tredoux *et al* (2005) 366-367; Walker (1989) 302; Ewing (1987) 51. According to Blackman (1986) 228-229, studies in psychology and sociology indicate that battered women may find it virtually impossible to leave abusive relationships as a result of psychological changes that follow after staying in an abusive relationship after a second episode of abuse. Blackman notes that there are three categories of change that occur in battered women motivating them to remain in an abusive relationship: (1) Most battered women experience psychological changes inducing them to believe that they are unable to control their fate and that they are unable to put an end to the abuse. They may become depressed and "learned helplessness" may ensue. (2) Battered women often show "high tolerance for cognitive inconsistency" in that they express two ideas that appear to be inconsistent with one another by for example claiming that the abuser was violent while drunk, but later recall an episode during which he was not drunk but was nevertheless abusive. (3) Battered women often experience a sense that alternatives are not available to them. They experience an inability to stop the violence and believe there is no escape from the relationship. Expert evidence is therefore necessary to clarify these issues.

143 1986 *Louisiana LR* 985-986. See also Crocker "The meaning of equality for battered women who kill men in self-defense 1985 *Harvard Women's LJ* 127.

144 Ewing (1989) 53; Hudsmith 985; Mather 574-576. See also *State v Kelly* 478 A2d 364 (1984) 377 where the New Jersey Supreme Court stated the following in respect of expert evidence: "Experts point out that one of the common myths, apparently believed by most people, is that battered wives are free to leave. To some, this misconception is followed by the observation that the battered wife is masochistic, proven by her refusal to leave despite the severe beatings, to others, however, the fact that the battered wife stays on unquestionably suggests that the beatings could not have been too bad for it they had been, she certainly would have left. The experts could clear up these myths, by explaining that one of the common characteristics of a battered wife is her inability to leave despite such constant beatings, her 'learned helplessness', her lack of anywhere to go, her feeling that if she tried to leave, she would be subjected to even more merciless treatment, her belief in the omnipotence of her battering husband, and sometimes her hope that her husband will change his ways."

Potential uses of expert evidence pertaining to the Battered Woman Syndrome are the following:¹⁴⁵

- To introduce the court to a class of persons – battered women, and their profile. Experts can also explain the cycle of violence, learned helplessness and also coercive control and other dynamics in respect of battering relationships.
- To provide the trier of fact with an explanation as to why the mentality and personality make-up and behaviour of battered women differ from the lay person's perspective of how someone would react towards an abusive partner.
- To indicate to the court that the accused and the victim were involved in an abusive relationship.
- To explain why the woman remained in the abusive relationship.
- To refute popular myths and misconceptions concerning battered women, including:¹⁴⁶
 - The myth that these women are masochistic.

145 Johann and Osanka (1989) 159–160; Walker (1989) 322–327; Ludsin and Vetten 93; Ewing (1987) 52–60; Roberts “Between the heat of passion and cold blood: Battered woman's syndrome as an excuse for self-defense in non-confrontational homicides” 2003 *Law and Psychology R* 149 151 who quotes an extract from *Ex parte Haney* 603 50.2d 412 (Ala 1992) 414 where the court stated: “Expert testimony on the battered woman syndrome would help dispel the ordinary lay person's perception that a woman in a battering relationship is free to leave at any time. The expert evidence would counter any “common sense” conclusions by the jury that if the beatings were really that bad the woman would have left her husband earlier. Popular misconceptions about battered women would be put to rest, including the beliefs that the women are masochistic and enjoy the beatings and that they intentionally provoke their husbands into fits of rage.”

146 See also *S v Engelbrecht* (*supra* fn 6) para 29 where the court per Satchwell J proceeded to summarise the principles upon which expert testimony should properly be admitted in cases where battered woman syndrome evidence is advanced: Expert testimony is admissible to assist the fact-finder in drawing inferences in areas where the expert has relevant knowledge or experience beyond that of the lay person; there are stereotypes, for instance, that battered women are not really beaten as badly as they claim otherwise they would have left the relationship, alternatively, that women enjoy being beaten because they have a masochist strain in them – which stereotypes may adversely affect consideration of a battered woman's claim to have acted in self-defence in killing her mate and expert evidence can assist in dispelling these myths; expert testimony relating to the ability of an accused to perceive danger from her mate may go to the issue of whether she “reasonably apprehended” death or grievous bodily harm on a particular occasion; expert testimony pertaining to why an accused remained in the battering relationship may be relevant in assessing the nature and extent of the alleged abuse; and by providing an explanation as to why an accused did not flee when she perceived her life to be in danger, expert testimony may also assist in assessing the reasonableness of her belief that killing her batterer was the only way to save her own life. See also para 28 where Satchwell J held that expert evidence on both the social context of domestic violence and on the specific effects of abuse on the psyche of an abused woman who kills, is essential. Expert evidence can assist the court to understand the unequal power relations in an abusive relationship which impact on the woman's ability to leave and the manner in which she resorts to violence; why abused women often do not leave the abusive relationship; and the process leading to the point at which she becomes psychologically unable to adjust to and accommodate the ever-present danger of abuse. Such evidence is necessary to refute widely recognized myths and misconceptions concerning battered women that would interfere with judge or juror ability to assess the woman's actions fairly.

- The myth that these women stay with their abusers because they enjoy beatings.
- The myth that these women could freely leave these abusive relationships if they really wanted to.¹⁴⁷
- To provide the court with an opinion of the accused's state of mind at the time of the commission of the crime.
- To rebut the implication of premeditation or planning.
- To support a defence of not guilty as a result of mental illness or mental defect.
- Battered Woman Syndrome evidence can be used in support of diminished responsibility and mitigation of sentence.

The function of expert evidence on battered women is to provide the court with an alternative perspective or social framework for understanding the particular woman's beliefs and actions. Expert evidence will also attempt to dispel any myths or misconceptions that the court may have as to the psycho-social dynamics and consequences of abuse.¹⁴⁸ In most instances the expert will explain the general research findings regarding battered women and also provide a clinical opinion that the particular accused displays signs of the syndrome or suffers from the syndrome.¹⁴⁹

The main obstacle in respect of expert evidence in support of the defence of non-pathological criminal incapacity is that it is not compulsory to advance expert evidence in support of this defence. Women who kill their abusers often do not do so in the midst of immediate confrontation. It is often only after a specific incident when a woman resorts to deadly force. Reliance on the defence of non-pathological criminal incapacity becomes difficult as a result of the non-confrontational killing of the abusive partner. The latter has resulted in abused women relying on alternative defences, often unsuccessfully.

It is important to keep in mind that expert evidence should not only be advanced in order to exonerate the accused, but also in support of a possible mitigation of sentence and also by the prosecution in order to refute the defence.

Ludsin and Vetten indicate that there are four main types of evidence that should be advanced on behalf of a battered woman who killed her abusive

147 See also *Lavallee v The Queen* (1990) SCR 85 (55cc) (3d) 97 (SCC) where the Canadian Supreme Court per Wilson J held: "Expert evidence on the psychological effect of battering on wives and common law partners must, it seems to me, be both relevant and necessary in the context of the present case. How can the mental state of the appellant be appreciated without it? The average member of the public (or the jury) can be forgiven for asking; why should a woman put up with this kind of treatment? Why would she continue to live with such a man? How could she love a partner who beat her to the point of requiring hospitalisation? We would expect the woman to pack her bags and go. Where is her self-respect? Why does she not cut loose and make a new life for herself? Such is the reaction of the average person confronted with the so-called battered wife syndrome. We need help to understand it and help is available from trained professionals."

148 Schuller and Vidmar 277. See also Monahan and Walker *Social science and law: Cases and materials* (1990) 277; Schneider "Describing and changing women's self defense work and the problem of expert testimony on battering" 1986 *Women's Rights Law Reporter* 195–222.

149 Schuller and Vidmar 277.

partner:¹⁵⁰ Evidence pertaining to the history and pattern of abuse that the battered woman endured during the relationship with the abuser; evidence pertaining to other violent acts of the abuser that the accused was aware of; social context evidence; and evidence in respect of other violent acts of abuse performed against the accused.

The history and pattern of abuse between the battered woman and the deceased are important aspects in support of a defence raised by a battered woman to a charge of murder or in support of mitigation of sentence.¹⁵¹ This history and pattern of abuse are important factors in explaining the battered woman's mental state at the time of killing her abusive husband or partner.¹⁵² The history and pattern of abuse may also be used to assist in explaining the cumulative effect of fear, stress and/or provocation that induced non-pathological criminal incapacity or diminished criminal capacity.¹⁵³

A history of the deceased's violence against others also assists in explaining the woman's reasonable fear of death or serious bodily harm that could be caused by the abuser.¹⁵⁴ Social-context evidence advanced by a battered woman who kills her abuser either in support of a defence or in support of mitigation, can be divided into two categories:¹⁵⁵

The first category explains how women are treated by the government, courts, family members and society in general. This type of evidence accounts for a woman's limited options for escaping the abusive relationship. This evidence elucidates the woman's frame of mind and attempts to place the court within the frame of mind of the battered woman. This evidence could also substantiate the credibility of the accused who claims non-pathological criminal incapacity or diminished criminal capacity.¹⁵⁶

The second category of social context evidence relates to psychosocial evidence. This type of evidence specifically pertains to the psychological effects of abuse on women.

Ludsin and Vetten¹⁵⁷ state that evidence of the psychology of batterers promotes the reasonableness of the effects of abuse on women; whilst evidence of social judgment explains that which motivates abused women to kill. Battered women who kill their abusers need to provide expert evidence pertaining to the psychological effects of abuse on women in general and also with specific reference to

150 187.

151 *Ibid.*

152 *Idem* 189; Schuller and Vidmar 276 state: "The violence that battered women faces is continual and is at the hands of an intimate partner rather than a stranger. Furthermore, the woman is generally not on equal physical grounds with the batterer. As a result, when she strikes back, her actions cannot be the same as a fight between 'two equals', and usually this is reflected in the circumstances surrounding the killing." See also Roberts (2005) 143–144, Ewing (1989) 52–54.

153 Ludsin and Vetten 189. See also Dobash and Dobash *Violence against wives* (1979) 6 who state: "When the man dies, it is rarely the final act in a relationship in which she has repeatedly beaten him. Instead, it is often an act of self-defence or a reaction to a history of the man's repeated attacks." See also *idem* 31–74; Browne (1987) 109–130; Mather 547–555.

154 Ludsin and Vetten 190.

155 *Ibid.* See also Ludsin fn 17 *supra*.

156 Ludsin and Vetten 192.

157 192–193; Mather 550–554.

the accused in order to establish the factual foundation for the defence of non-pathological criminal incapacity.¹⁵⁸

Within the context of non-pathological criminal incapacity, the effects of abuse on the woman are vital in order to provide clarity as to why she lost control at the time of the act.¹⁵⁹ Evidence of prior acts of violence committed against the battered woman should also be tendered in expert testimony pertaining to the psychological effects of abuse on the particular woman, also with specific reference to the defence of non-pathological criminal incapacity.¹⁶⁰

6 EXPERT EVIDENCE NECESSARY IN SUPPORT OF A CLAIM OF DIMINISHED CRIMINAL CAPACITY IN RESPECT OF SENTENCING

A battered woman could also, as an alternative to the defence of non-pathological criminal incapacity due to provocation and emotional stress, rely on diminished criminal responsibility as espoused in section 78(7) of the Criminal Procedure Act, in which event the diminished criminal capacity will only have a bearing on mitigation of sentence.¹⁶¹ In the recent decision of *S v Mnisi*,¹⁶² the appellant's sentence of eight years imprisonment was reduced to five years on appeal due to his diminished criminal responsibility at the time of the commission of the offence. In the latter decision the appellant pleaded guilty in terms of section 112(2) of the Criminal Procedure Act, but claimed that he had acted with diminished criminal responsibility. On appeal it was argued, *inter alia*, that the trial court had not afforded sufficient weight to the fact that the appellant had acted with diminished responsibility. The majority of the court held that the appellant had established a sufficient factual foundation to support a finding of diminished responsibility when he committed the offence. The court specifically noted that the appellant did not seek to rely on the defence of temporary non-pathological criminal incapacity, but rather upon diminished responsibility which is not a defence but is relevant to the question of sentence.¹⁶³ It was also held that the question whether an accused acted with diminished responsibility must be assessed in light of all the evidence, expert or otherwise.¹⁶⁴ It was further held that there is no obligation upon an accused to adduce expert evidence and that the *ipse dixit* of an accused may suffice, provided that a proper factual foundation is laid.

158 *Ibid.*

159 Ludsin and Vetten 193. They also state that expert evidence can counter the presumption of goal-directed behaviour that normally leads to the suggestion of full criminal capacity. See also Lenkevich (1999) 318.

160 Ludsin and Vetten 196. They note that when the defence of non-pathological criminal incapacity is raised, legal practitioners need to ascertain the following: Why the woman killed her husband; whether there was a triggering event; how the woman reacted before, during and after the killing; when the incapacity started; and whether she regained capacity at any point between when she first lost criminal capacity and when she killed.

161 See Snyman (2008) 176; s 78(7) of the Criminal Procedure Act. See also *S v Campher* (*supra* fn 6) 964C–E; *S v Smith* (*supra* fn 6) 135b–e; *S v Shapiro* (*supra* fn 19) 123c–f; *S v Ingram* (*supra* fn 19) 8d–e.

162 *S v Mnisi* 2009 2 SACR 227 (SCA). See also Carstens 2010 *De Jure* 388–394.

163 Para 4. This *dictum* by the majority of the court could be construed as being an indication that the defence of non-pathological criminal incapacity is still recognised as a valid defence.

164 Para 5.

Carstens notes in respect of the former decision, and this view is supported by the author, that expert evidence is now essential to sustain not only the defences of sane automatism and/or non-pathological criminal incapacity, but also in respect of diminished incapacity/responsibility.¹⁶⁵ Carstens correctly notes: “As the diminished responsibility is a question of degree and may be a precursor for a successful defence of at least non-pathological automatism due to provocation . . . the accused’s diminished ability also needs to be assessed by experts.”¹⁶⁶ Carstens also notes that it is a pity that the appellant did not rely on the defence of non-pathological criminal incapacity as it could have resulted in a reassessment of the much debated *Eadie* controversy.¹⁶⁷

It is clear that a battered woman may also rely on diminished criminal capacity in support of mitigation of sentence, and it is submitted that also in such instance expert evidence will play a pivotal role to explain the dynamics of the abusive relationship and those aspects which diminished the battered woman’s criminal capacity. As Carstens notes, it would be advisable to note a plea of not guilty in such instance and to order the state to prove the charge against the accused beyond reasonable doubt.¹⁶⁸ In such an instance both the prosecution and the defence should produce expert evidence pertaining to the accused’s criminal capacity at the time of the commission of the offence and will empower the court to determine whether such criminal capacity was indeed diminished and to what extent.¹⁶⁹

7 CONCLUSION

Abuse within intimate relationships is an inescapable reality of society. In this article the pivotal and essential role of expert evidence in respect of abused women was illustrated and contextualized against the backdrop of a selection of psycho-social dynamics of abuse and the various explanations advanced as to why women endure abuse rather than escape the abusive relationship. The latter was canvassed against the backdrop of the defence of non-pathological criminal incapacity which constitutes one of the defences available to a battered woman. The merits of the defence of non-pathological criminal incapacity were also addressed within the context of the highly-debated and controversial decision in *Eadie*. The complexity of the issue and also the fact that expert evidence in cases of non-pathological criminal incapacity fulfils a central function, contrary to the traditional approach in terms of which expert evidence does not fulfil a central function, was elucidated. The presentation of expert evidence will ensure the accused’s right to adduce and challenge evidence and also lead to a fair and a just trial provided that the state, in turn, presents experts to challenge the evidence advanced on behalf of the accused. In cases where non-pathological criminal incapacity is raised by abused women, the full purport of their constitutional rights infringed during the abuse will receive due recognition in line with a constitutional state founded on the values of human dignity, equality and freedom.

165 Para 5.

166 Carstens 2010 *De Jure* 393.

167 *Idem* 394.

168 *Ibid.*

169 *Ibid.*

The legal regulation of access and benefit-sharing with regard to human genetic resources in South Africa

M Nöthling Slabbert

BA BA(Hons)MA DLitt LLB LLD

Professor, Department of Jurisprudence, University of South Africa

OPSOMMING

Regsregulering van toegang tot en deling van voordele wat spruit uit menslike genoomnavorsing

Die artikel bespreek krities die regsregulering van die toegang tot en die deling van voordele wat spruit uit menslike genoomnavorsing. Huidige wetgewing in Suid-Afrika maak nie voorsiening vir die regulering van toegang tot en die deling van voordele voortspruitend uit menslike genoomnavorsing nie. Die artikel ondersoek internasionale regsinstrumente wat die toegang tot en deling van voordele van genoomnavorsing reguleer; potensiële probleme met die toepassing van hierdie aspekte, asook moontlike toepaslike regsmodelle vir die Suid-Afrikaanse konteks in die lig van die Suider-Afrikaanse Menslike Genoomprogram wat onlangs in die lewe geroep is. Kwessies soos toegang tot en deling van voordele van menslike genoomnavorsing is tans twee kritiese aspekte waarvoor duidelikheid in die Suid-Afrikaanse regs konteks vereis word, veral teen die agtergrond van Suider-Afrika se ryk mens-, plant- en diergenetiese en biodiverse erfenis.

1 INTRODUCTION

Indigenous populations in southern Africa are generally perceived to be the most genetically divergent on the planet and hence a very valuable source of genetic information. This potential wealth of knowledge, unfortunately, does not necessarily mean that southern Africa will automatically reap the benefits of knowledge derived from its own genetic resources, which may include financial gain in the form of intellectual property protection (eg patents) and commercialisation of new health solutions.¹

Biodiversity legislation in South Africa is weak in terms of the recognition of the customary rights of rural communities vis-à-vis the rights of the State to regulate access and use of the country's biological resources, including the enforceability of the rights of rural communities and the proper inclusion of prior informed consent.² This position is exacerbated by the reality that although

1 Jakobsson *et al* "Genotype, haplotype and copy-number variation in world-wide human populations" 2008 *Nature* 998–1003; Pepper "The human genome and molecular medicine: Promises and pitfalls" 2010 (100)11 *SAMJ* 722–723 723.

2 Marcelin, for example, argues that the National Environmental Management: Biodiversity Act of 10 of 2004 does not include adequate provisions reflecting customary laws and introducing prior informed consent. See Marcelin "Biodiversity regulatory options:

continued on next page

benefit-sharing has been established as a principle of international law in the context of biodiversity, a legal framework relevant to benefit-sharing in respect of *human* genetic resources is lacking. Access to and benefit-sharing for human biological resources (eg blood and DNA-samples) are presently not regulated through a clear international legal framework, such as the Convention on Biological Diversity (CBD) and the recently adopted Nagoya Protocol, both of which only regulate access to plants, animals, micro-organisms and traditional knowledge.³ In South Africa, access to human genetic resources and benefit-sharing is uncertain in view of the fact that there are presently no national guidelines governing genomic research and its legal or ethical ramifications that succeed in balancing the protection of human genetic information and the promotion of international collaboration that may increase the development of local scientific capacity.⁴ Human genetic material may leave South Africa's borders virtually undetected and undocumented, as has already happened in some instances.⁵

A recent article⁶ investigating the notions of genomic sovereignty and benefit-sharing in South Africa, following a study that was published in *Nature*⁷ involving the genetic sequencing of the genomes of Archbishop Desmond Tutu and a number of Khoisan people, questioned inter alia the long and short-term benefits that this predominantly foreign study would hold for southern Africa and its peoples. This concern is justifiable against the background of other controversial cases of biomedical research in Africa such as the patent application on African lactose tolerance genes, as well as the patent on the recently discovered retrovirus, Human T-Lymphotropic virus (HTLV), found in

Involvement of rural communities in decision-making processes in South Africa" 2006 (8)6 *J of World Intellectual Property* 809–824. Rural communities are often not involved in the decision-making process through which the research or plant collection permits are issued (or refused) to prospective applicants, and thus they are not legally empowered to control access and to minimise the misappropriation and overexploitation of resources in their territory.

3 See Alvarez Núñez "Intellectual property and the protection of traditional knowledge, genetic resources and folklore: The Peruvian experience" 2008 (12) *Max Planck Yearbook of United Nations Law* 487 for a discussion of the protection of traditional knowledge, genetic resources and folklore, with reference to benefit-sharing. The Nagoya Protocol on Access to Genetic Resources and the Fair and Equitable Sharing of Benefits Arising from their Utilization to the Convention on Biological Diversity is an international agreement regulating the sharing of benefits arising from the utilisation of genetic resources in a fair and equitable way, including by appropriate access to genetic resources and by appropriate transfer of relevant technologies. It was adopted by the Conference of the Parties to the Convention on Biological Diversity at its tenth meeting on 29 October 2010 in Nagoya, Japan. The Nagoya Protocol has opened for signature by Parties to the Convention from 2 February 2011 until 1 February 2012. The fair and equitable sharing of the benefits arising out of the utilisation of genetic resources is one of the three objectives of the Convention on Biological Diversity.

4 See in general Slabbert & Pepper "'A room of our own?' Legal *lacunae* regarding genomic sovereignty in South Africa" 2010 *THRHR* 432–450.

5 See eg McGown *Out of Africa: Mysteries of access and benefit sharing* (2006) and *Pirating African heritage* (2009).

6 Slabbert & Pepper fn 4 above.

7 (2010) 463 857.

African blood samples.⁸ These studies were criticised for problems relating to informed consent and the consideration of benefit-sharing. Documented elsewhere in the world is the misuse of biological samples by researchers in respect of the indigenous populations of the Havasupai Indians⁹ and the Nuu-chah-nulth.¹⁰

It is a sad reality that developing countries are not participating fully in genomics research (eg as a result of limited human and financial resources and lack of technical capacity, including bioinformatics).¹¹ The question as to how these countries may benefit from discoveries in this rapidly developing field, particularly those that are done on genetic material gathered from within their borders, is beginning to receive more attention.¹² This question is a very important one, considering the increasing emphasis on the concept of genomic sovereignty, which can be described as “representing a nation’s ability to capture the value of its investments in the field of genomic medicine”.¹³ It refers to the capacity of a people, a country or nation to own, to control both access to and use of, samples, data and knowledge concerning human genes.¹⁴ The notion of genomic sovereignty appears to contradict the 1997 statement of the United Nations Educational, Scientific and Cultural Organisation (UNESCO) which

8 Hammond & Mayet *Genes from Africa: The colonization of human DNA* (February 2009), available at <http://www.biosafetyafrica.org.za>. See also *Pirating African heritage – The pilaging continues* (2009) at <http://www.biosafetyafrica.org.za> (accessed 12 April 2011).

9 In November 2008, Havasupai Indians were granted permission by the Arizona Court of Appeals to proceed with legal action related to the misuse of biological samples taken by the Arizona State University and the University of Arizona. In the early 1990s, more than 200 genetic samples were consensually loaned for diabetes research. Years after the findings of this study were published the samples were used without the consent of the Havasupai to investigate schizophrenia, inbreeding and population migrations. For more detail, see Davenport “Court revises Arizona tribe’s lawsuit over research” available at <http://bit.ly/iXXayE> (accessed 28 February 2010).

10 More than 25 years ago, approximately 800 genetic samples were taken from the Nuu-chah-nulth tribe by researchers at the University of British Columbia. The Nuu-chah-nulth gave consent that the samples be used to investigate rheumatoid arthritis, which affects the tribe significantly. However, it appeared later that the samples were used and shared with other researchers across the world, without their consent, for a variety of different studies. The tribe never benefited in any way from any of these later studies. See Dalton “Tribe blasts ‘exploitation’ of blood samples” 2002 *Nature* 420. See also Pellekaan “Genetic research: What does this mean for Indigenous Australian communities?” March 2002 *Australian Aboriginal Studies* 65–75.

11 Coloma & Harris “Molecular genomic approaches to infectious diseases in resource-limited settings” 2009 (6)10 *PLoS Medicine* 1.

12 See Harmon “DNA gatherers hit snag: Tribes don’t trust them” *New York Times* 10 Dec 2006. Some of these DNA studies may also contradict specific indigenous cultural histories, potentially thwarting legal sovereignty arguments and other legal claims. DNA studies, in conjunction with other evidence, may corroborate claims relating to land, eg by suggesting, based on migration patterns, that one ethnic group occupied a given area prior to others.

13 Hardy *et al* “South Africa: From species cradle to genomic applications” (Oct 2008) 9 *Nature Review Genetics* S20.

14 For a more detailed discussion on the legal aspects relating to genomic sovereignty in South Africa, see Slabbert & Pepper fn 4 above.

declares the human genome “the heritage of humanity”.¹⁵ This statement in effect labels the human genome as an asset *extra commercium*. The danger associated with the 1997 UNESCO view is that it may lead to “bio-colonialism” or “genetic piracy” of genetic samples in developing countries lacking the scientific capacity or resources to conduct the research themselves.¹⁶ This suggests a differentiation between human and non-human genetic resources, with the first-mentioned seen as part of the common heritage of humanity but the latter as belonging to the sovereign right of states.¹⁷ This will be explored in more detail below.

The need for a consideration of the issues relevant to the access to human genetic resources, as well as benefit-sharing is extremely relevant in the present South African context in view of the fact that the southern African Genome Programme has recently been launched.¹⁸ The initial aim of this Programme is to sequence whole genomes of individuals of African origin. The medium-to-long term objectives are to (1) determine genetic risk factors and pathogenetic mechanisms of diseases that are important to the southern African population (eg infectious diseases and diseases of lifestyle); and (2) to build bioinformatic capacity across the country in order to be able to analyse data generated locally and also to access and analyse data published in the public domain that could have an impact on our local populations.

Access and benefit-sharing of human genetic resources are two aspects that will need to be carefully examined in the execution of the above Programme.

Against this background, the purpose of this article is to examine the concepts of access to human genetic material and benefit-sharing in South Africa, with reference to relevant international documents that may inform legislators in this field and also drafters of benefit-sharing agreements. The article will also consider a few benefit-sharing models. Excluded from the scope of this article is a discussion of informed consent relating to the donation of biological samples prior to the conclusion of benefit-sharing agreements, as well as relevant human rights issues. Aspects regarding commercialisation and intellectual property rights in respect of information derived from genetic sequencing will also not be specifically addressed in this article. These issues impact on the translation of genetic discoveries into health benefits and require detailed consideration in their own right.

2 ACCESS AND BENEFIT-SHARING

2.1 What are access and benefit-sharing exactly?

The idea of benefit-sharing has its origin in the perception of injustice that may emerge from inequalities of power between the global medical industry and

15 See UNESCO, Universal Declaration on the Human Genome and Human Rights, General Conference Res 29 C/Res16, reprinted in Records of the General Conference, UNESCO, 29th Sess, 29 C/Resolution 19, 41 (1997) (adopted by the UN General Assembly, GA res 152, UN GAOR, 53rd Sess, UN Doc A/RES/53/152 (1999)).

16 See eg in general Cohen, Illingworth & Schueklenk (eds) *The power of pills: Social, ethical and legal issues in drug development, marketing and pricing* (2006) 203–215.

17 As encapsulated by a 3 of the CBD.

18 Application submitted to the Department of Science and Technology in December 2009.

the resource-rich, yet less developed countries. As a mechanism to restore equilibrium to the asymmetry between researcher and participant in the spirit of deontological considerations of fairness, it has become a specific principle of research governance consistent with social justice.¹⁹ Although medico-ethical guidelines and regulations generally emphasise the responsibilities of sponsors and researchers,²⁰ it is clear that the increasing influence and power of non-state actors in the field of medical research have brought more responsibilities for the last-mentioned. That benefit-sharing is based on the language of human rights is clear from the *Statement on benefit sharing* of the HUGO Ethics Committee. This Committee suggests that companies sponsoring research should set aside 1-3% of the annual net profit to support health care infrastructures and/or humanitarian efforts.²¹

Access and benefit sharing have traditionally been relevant in two areas: biobanks (DNA and the data derived from the sequencing thereof) for the purpose of pharmacogenetics and pharmacogenomics as well as population genomics research, and the use of traditional knowledge from indigenous communities²² to develop new products (eg in the pharmaceutical industry, involving non-human genetic resources).

It is generally accepted that benefit-sharing involves not only the distribution of benefits but also the burdens arising from research and development activities in human genetics. This is the case as results following research may include a wide array of potential benefits and burdens, ranging from improved drugs and therapies, to preventive and personalised medicine, with neither benefits, nor burdens, necessarily linked to the health care sector.²³

Although access and benefit-sharing are normally referred to as one concept (with the acronym ABS), access itself refers to “the ability of individuals to

19 Emerson *et al* “Access and use of human tissues from the developing world: Ethical challenges and a way forward using a tissue trust” 2001 12(2) *BMC Medical Ethics*, available at <http://www.biomedcentral.com/1472-6939/12/2> (accessed 22 April 2011).

20 Eg guidelines 10 and 21 of the Council for International Organizations of Medical Sciences (CIOMS): *Guidelines for biomedical research involving human subjects* (available at <http://bit.ly/jPcrqS>, accessed 23 April 2011); World Medical Association, *Helsinki Declaration*, principle 30, available at <http://bit.ly/iJF4PO> (accessed 24 April 2011); European Group on Ethics in Science and New Technologies, *Ethical aspects of clinical research in developing countries*, available at <http://bit.ly/kMlgDS> (accessed 24 April 2011).

21 Recommendation 6; HUGO Ethics Committee’s *Statement on Benefit Sharing*, available at <http://bit.ly/ky1Vru> (accessed 22 April 2011).

22 An example arising in the Southern African context in respect of non-human genetic resources is the benefit-sharing agreement signed in 2003 between the San people on the one hand, and the CSIR and Phytopharm regarding the San’s traditional plant knowledge of the appetite-suppressing properties of the *Hoodia* plant that was developed into treatment for obesity. See Wynberg “Rhetoric, realism and benefit sharing” 2004 (7) *J of World Intellectual Property* 854–876. This agreement was only signed after legal proceedings were initiated by the San people.

23 Simm *Benefit sharing frameworks: Justifications for and against benefit sharing in human genetic research* (Report for GenBenefit 2007), available at <http://bit.ly/kkje34> (accessed 22 April 2011).

acquire, exchange, or use genetic resources for a multitude of purposes, not necessarily limited to commercial application".²⁴

The next logical question is what exactly are or could be considered to be benefits? A clarification of the potential categories of benefits is necessary. The HUGO Ethics Committee, for example, defines a benefit in its *Statement on benefit sharing* as follows:

"A benefit is a good that contributes to the well-being of an individual and/or a given community (e.g. by region, tribe, disease group). Benefits transcend avoidance of harm (non-maleficence) in so far as they promote the welfare of an individual and/or of a community. Thus, a benefit is not identical with profit in the monetary or economic sense. Determining a benefit depends on needs, values, priorities and cultural expectations."²⁵

Access and benefit-sharing may take the form of access to medical care, products, technologies or new drugs; financial payment; technology transfer; research cooperation; capacity building; the wide dissemination of knowledge; joint ventures; reimbursement of costs; use of a portion of royalties for humanitarian purposes; and the accessibility of research results to the biobank or database. Benefits may not be immediate or tangible, as some studies are conducted over extended periods, with benefits accruing in the future and not for the immediate direct benefit of the participants themselves.²⁶ Even if research does not lead to any products, valuable information may be gained which can guide further research. Also relevant is the question of how and when benefits must be defined: are they determined before the research commences; can they be renegotiated during the research process; can a community or participants define the desired benefit (eg a benefit totally unrelated to the outcome of the research)?

Benefits should also be appropriate and not, for example, refer to the supply of DNA samples in exchange for toothbrushes.²⁷ The specific benefit should accord with the ability of the community to utilise it, eg technology transfer²⁸ will not be useful to a community that lacks trained personnel, whereas a medical benefit may require a basic infrastructure. Benefits should also be proportional to the significance of the contribution. Provisions relating to a share in the intellectual property rights are more common in the regulations of more advanced developed countries that manage their own genetic research programmes, such as India, South Africa and Brazil.

24 Ghose *Benefit sharing of genetic resources* (Gene Campaign: Research project on Protection of Indigenous Knowledge of Biodiversity, briefing paper 1, August 2004), available at <http://ranaghose.com/research/gcabs.pdf> (accessed 26 April 2011).

25 Available at <http://www.eubios.info/BENSHARE.htm> (accessed 22 April 2011).

26 Knoppers, Abdul-Rahman and Bédard "Genomic databases and international collaboration" 2007 (18) *King's LJ* 291–312 304.

27 Emerson *et al* 2.

28 Technology transfer is emphasised in a number of international law and policy guidelines, a 7 of the TRIPS agreement (*Agreement on trade-related aspects of intellectual property rights*; Annex 1C of the Marrakesh Agreement establishing the World Trade Organization, signed in Marrakesh, Morocco on 15 April 1994, text available at <http://bit.ly/j2UHkr>, accessed 24 April 2011) emphasises that technology transfer should take place in a manner conducive to social and economic welfare and to a balance of rights and responsibilities.

The following table illustrates some of the types of benefits that can be identified:²⁹

	Health benefits	Commercial/economic benefits	Scientific benefits
Individual level	Designer drugs/ individual aspects of personalised medicine	Profits to investors	Non-instrumental Knowledge Development of science and gaining of new information
Communal level	Relief to disease- affected populations	Non-medical benefits to communities; capability-building	As above
National, state level	Efficient health care services; policy planning	Development of biotechnology; new job opportunities; capability building	As above
Global level	Eradication of diseases		As above

Despite UNESCO's incorporation of the principle of benefit-sharing into the International Declaration on Human Genetic Data,³⁰ and although benefit-sharing appears to be a recurrent theme in international debates relating to the use of non-human genetic resources, the concept has also never been satisfactorily defined.³¹

Comparing some of the national definitions on benefit-sharing of parties to the CBD (relevant of course in respect of non-human genetic resources only), Schroeder³² comments on the "puzzling" nature of some of these definitions, which she describes as being either unclear or not definitions at all. One exception is the definition on a regional level by the African Union which describes benefit-sharing as the sharing of whatever accrues from the use of biological resources, community knowledge, technologies, innovations or practices.³³ As she points out, however, this definition fails in one respect,

²⁹ Simm 17.

³⁰ A 19(a) states that "[b]enefits resulting from the use of human genetic data, human proteomic data or biological samples collected for medical and scientific research should be shared with the society as a whole and the international community".

³¹ See Schroeder "Benefit sharing: It's time for a definition" 2007 (33) *J of Medical Ethics* 205. The origin of the term is the Convention on Biological Diversity.

³² Schroeder 206. One of the national definitions that he refers to is that of the UK, which defines benefit-sharing as the "sharing of benefits arising from the use, whether commercial or not, of genetic resources, and may include both monetary and non-monetary returns".

³³ See OAU, African Model Legislation for the Protection of the Rights of Local Communities, Farmers and Breeders, and for the Regulation of Access to Biological Resources, available at <http://bit.ly/kNsB4V> (accessed 19 April 2011). A 12 of this Model Legislation provides that a payment be made by those seeking collection of samples prior to the collection process itself.

namely that it does not specify that two parties have to participate in a legal benefit-sharing process, otherwise the arrangement is one of charitable giving. Flowing from this defect are two conclusions, namely that the definition suggests that benefits may be demanded from those in no way connected with the used resource; and that all uses of resources warrant benefit-sharing.

Moreover, the reference to *biological resources* instead of *genetic resources* should be carefully considered, as a biological resource – which is much broader than a genetic resource – would include, for example, derivatives and biochemical reactions which are excluded from a genetic resource. Using the terms interchangeably would extend the scope of application to include innovations, for example, resulting from the use of resources.³⁴

One of the clearest definitions of benefit-sharing emphasising the current imbalance between genetic research activities and the distribution of benefits among those contributing to the research, as well as the danger of undue inducement,³⁵ is the one suggested by Schroeder:

“Benefit sharing is the action of giving a portion of advantages/profits derived from the use of human genetic resources to the resource providers to achieve justice in exchange, with a particular emphasis on the clear provision of benefits to those who may lack reasonable access to resulting healthcare products and services without providing unethical inducements.”³⁶

2.2 Justification for benefit-sharing in the context of human genetic resources

Contrary to the wide agreement generally in respect of the normative justification for benefit-sharing in the context of non-human genetic resources, divergent views relating to benefit-sharing pertaining to human genetic resources exist.

2.2.1 Arguments against benefit-sharing

Two arguments against benefit-sharing can be distinguished, namely one that focuses on the necessity and supremacy of altruism in human genetic research, and another that views any profiting from the human genome as inherently immoral, as it is believed that the value of the human genome cannot be represented in monetary or similar trade-off terms.

In terms of the first-mentioned view, altruism should be the guiding principle for contributors to human genetic research.³⁷ The last-mentioned view, namely that genetic material is sacred, priceless or non-commercialisable,³⁸ flows from various understandings of humanist and religious accounts of human bodily in-

34 Schroeder 206.

35 In line with UNESCO’s Universal Declaration on Bioethics and Human Rights, which requires that benefit-sharing occurs and that improper inducements are avoided; adopted 19 October, 2005, text available at <http://bit.ly/k57fOG> (accessed 5 March 2010).

36 Schroeder 208. The line between unethical inducement and appropriate benefit-sharing is a fine one. Thambisetty argues that this should not be an obstacle to the development of appropriate benefit-sharing mechanisms, especially in respect of developing countries. See Thambisetty *Human genome patents and developing countries* Exeter College: University of Oxford, Study paper 10, prepared for the Commission on Intellectual Property Rights (2001) 35, available at <http://bit.ly/lqXdzr> (accessed 21 April 2011).

37 Berg “The ethics of benefit sharing” 2001 (59) *Clinical Genetics* 240–243.

38 See in general Resnik “The morality of human gene patents” 1997 (7)1 *Kennedy Institute of Ethics J* 43–61.

tegrity. In the context of human genetics and the huge profits to be gained from this field of research, it seems hypocritical to demand altruistic donations from participants in poor and developing countries, whilst industry-led researchers from developed countries reap all the benefit, especially if the purpose of the “genetic mining” in less developed countries is to produce products aimed solely at affluent populations of industrialised countries.³⁹ Also, promising a share of benefits to participants seems to contravene ethical principles that the body or human genome should not give rise to financial gain.⁴⁰ In addition, the ethical validity of consent given under the promise of benefits to be gained is also questionable.⁴¹ (The flipside of this is of course that although there is ethical objection to making a profit from genetic material, human DNA may in reality be considered as already commercialised, and as such, appropriate compensation for individuals in exchange for commercial access to their DNA would be in accordance with respect for human rights.)⁴²

2.2.2 Arguments in favour of benefit-sharing

The earliest applications of benefit-sharing originate from plant genomics and relate to agricultural resources, as find expression in the UN Convention on Biological Diversity. It is the agricultural framework that has furnished benefit-sharing with an argument founded on the notion of property, eg that the non-human genetic resources are owned in some sense, strengthened by the patenting of various plant and animal resources. The property/ownership argument in the context of human genetics, however, is more controversial, as ownership in this sense includes both the aspect of control over a resource or the capability to subject this resource to commercial transaction.⁴³ As mentioned above, the reference in the UNESCO Declaration on the Human Genome and Human Rights to the human genome as the “heritage of humanity” suggests the concept of common or shared property in the human genome in a symbolic sense.⁴⁴

By ethical and legal convention, no ownership rights over parts of the human body *ex vivo* exist, yet gene patents are regularly granted in respect of isolated gene sequences with known functions. Ownership over or control of human genetic resources is a deeply contested matter with a complex history, which will briefly be considered below.

There has recently been a gradual shift from viewing benefit-sharing not strictly in terms of the principles of medical ethics, emphasising compensation for risks taken, but also in terms of the principles of social justice and solidarity.⁴⁵ The notion of solidarity, common in the communitarian discourse, appears to be (a somewhat nebulous) one that attempts to promote the health interests universally, beyond the limited obligations of the traditional medical

39 Simm 8.

40 Thambisetty 6.

41 *Ibid.*

42 WHO: Genomics and World Health: Report of the Advisory Committee on Health Research (2002), available at <http://bit.ly/lpiMYg> (accessed 24 April 2011) 145.

43 Simm 10. See also Chadwick & Berg “Solidarity and equity: New ethical frameworks for genetic databases” 2001 (2)4 *Nature Review Genetics* 318–320.

44 See fn 15 above.

45 HUGO Statement on Benefit Sharing. See also Chadwick & Berg 318–321.

research setting. Universal benefit-sharing is ultimately set to respond to the call for global (distributive) justice that extends beyond compensatory or procedural justice.⁴⁶ The inherent danger of this aim is that benefit-sharing may be seen as an attempt to redress gross social inequities or create the kind of social upliftment of impoverished communities that rather belongs to the domain of government responsibility.⁴⁷ Unrealistic expectations should be avoided at all costs.

Many developing countries make benefit-sharing conditional to the authorisation of human genetic research. Two distinct advantages flowing from this are that individual participants will be less likely to be enticed into research with the promise of personal gain (with collective benefit dominating), and since the terms of the benefit-sharing in such instance will be contractual or statutory, national authorities will have greater control over the finer detail of such arrangements.⁴⁸ Ultimately, the aim of benefit-sharing is to build greater trust and reciprocity with participants, whilst at the same time countering potential exploitation through transparency and consultation.⁴⁹

The benefits or disadvantages of benefit-sharing cannot be properly assessed without a few brief remarks on the issue of sovereign control of the “building blocks of life”, considered next.⁵⁰

2 2 2 1 Control of genetic material

State sovereignty in plant agricultural resources falling within the borders of a recognised state is a well-established principle in international law, which is affirmed in article 15 of the CBD.⁵¹ Having said this, state sovereignty in respect of genetic resources within its territory is not absolute, as a tension may exist between the claims of indigenous communities within a sovereign state to their indigenous genetic resources (plant and human), recognised in the United Nations Declaration on the Rights of Indigenous Peoples,⁵² and those of the state

46 The HUGO Statement on Benefit Sharing refers to all three forms of justice, with specific emphasis on distributive justice. See also, *Who owns science? The Manchester Manifesto* (University of Manchester, Institute for Science, Ethics in Innovation), available at www.manchester.ac.uk/isei (accessed 26 April 2011).

47 Schuklenk & Kleinsmidt 8.

48 Thambisetty 35.

49 Knoppers “Population genetics and benefit-sharing” 2000 (3) *Community Genetics* 212–214.

50 Safrin “Hyperownership in a time of biotechnological promise: The international conflict to control the building blocks of life” 2004 (98) *Am J of Int L* 641–685.

51 Mgbeoji “Beyond rhetoric: State sovereignty, common concern, and the inapplicability of the common heritage concept to plant genetic resources” 2003 (16) *Leiden J of Int L* 821–826.

52 A 31 of the Declaration states that: “Indigenous peoples have the right to maintain, control, protect and develop their cultural heritage, traditional knowledge and traditional cultural expressions, as well as the manifestations of their sciences, technologies and cultures, including human and genetic resources, seeds, medicines, knowledge of the properties of fauna and flora, oral traditions, literatures, designs, sports and traditional games and visual and performing arts. They also have the right to maintain, control, protect and develop their intellectual property over such cultural heritage, traditional knowledge, and traditional cultural expressions.”

over the genetic resources within its borders,⁵³ in other words, multiple ownership interests may compete.

Traditionally, genetic material belonged to the global commons or open system, with no one exclusively owning this material and countries freely sharing it. Today, however, exclusive ownership of and restrictions on the sharing of genetic material appears to be the international norm.⁵⁴

Safrin⁵⁵ refers to the “legal enclosure of genetic material” which she ascribes to two developments: firstly the patenting of genetic material by developed countries, and secondly, in response to the privatisation of genes through the patent system, the extension of sovereignty over genetic resources by developing countries (who house most of the world’s raw genetic material). Developing countries are increasingly asserting state sovereignty rights or national governmental control over a wide range of genetic resources within their borders,⁵⁶ leading to what Safrin⁵⁷ describes as a “corrosive interplay” between the patent system and the sovereign-based system of developing states, which ultimately culminates in a system of “hyperownership” of genetic material.

The development of a system of extensive sovereign control or ownership of genetic research is said to (1) create the situation where multiple persons or entities hold rights of exclusion to a given resource; (2) pose a potential threat to the autonomy and liberty of persons and indigenous communities whose property contains such material; and (3) is based on an approach in international law which has led to unenforceable regimes regarding the regulation of access to genetic resources.⁵⁸ It has been argued that, collectively, this development has overreached in permitting or asserting legal ownership over genetic material.⁵⁹ Another consequence not expressly articulated by Safrin⁶⁰ is that this development has indeed led to unnecessary legal wrangling and overregulation which have the effect of curtailing human genetic research in developing countries.

At present, laws restricting access to genetic material with the purpose of obtain remuneration for a relevant state exclude human genetic material from their ambit.⁶¹ Globally, issues of access to *human* genetic resources at present appear to focus less on the question regarding sovereignty or ownership but more on the issue of benefit-sharing and the fair and equitable sharing of benefits between developed and developing countries. There is, however, a growing trend toward viewing human genetic material as a national resource or national patrimony,⁶² captured by the concept of human genomic sovereignty, referred to

53 See also the Joint Statement of the Indigenous Peoples Council on Biocolonialism, Call of the Earth, and the International Indian Treaty Council of 6 December 2006, available at <http://www.ipcb.org/> (accessed 24 April 2011).

54 Safrin 641.

55 *Ibid.*

56 Since 1993, close to 50 countries have passed legislation that restrict access to the genetic resources within their borders. See Safrin 641.

57 *Ibid.*

58 Safrin 641–642.

59 *Ibid.*

60 *Ibid.*

61 South Africa’s Biodiversity Act is one such example.

62 Slabbert & Pepper fn 4 above.

above, that refers to the capacity of a people, a country or nation to own, to control both access to and use of, samples, data and knowledge concerning human genes.⁶³ The traditional distinction in legal regimes between human and non-human genetic resources is slowly being eroded, with increased international lobbying calling for the CBD to govern access to human genetic material as well.⁶⁴ In 2002, an advisory committee to the WHO on genomics and world health, for example, referred in a report to the approach of the CBD relating to benefit sharing as a potential guide for such sharing in the human genetics area.⁶⁵

3 THE LEGAL FRAMEWORK FOR ACCESS TO AND BENEFIT-SHARING OF HUMAN GENETIC RESOURCES

3.1 The international context⁶⁶

Section 39(1)(b) of the Constitution instructs that when the Bill of Rights is interpreted, a court, tribunal or forum must consider international law. In the absence of clear domestic legislation regarding the regulation of access and benefit-sharing in the context of human genetic research, relevant international human rights documents may be instructive. The discussion will next turn to international law relevant to access and benefit-sharing, before turning to the position under South African law.

At the international level, the last ten years have seen the development of guidelines and statements of principle by various international bodies in respect of genomic research. Four different types of bodies relevant in this context may be distinguished, eg those representing all countries, such as the UN in the form of UNESCO; the Council of Europe representing countries within Europe and others who have signed its conventions; various international scientific

63 One of the world's most important and influential regimes restricting access to raw genetic material is the regional Common System on Access to Genetic Resources (Common System), promulgated by the Andean Pact nations of Bolivia, Colombia, Ecuador, Peru, and Venezuela. These five jurisdictions own the largest proportion of the world's biological diversity. Under the Common System, ownership of raw genetic material and derivatives of such genetic resources, such as molecules, vests with the relevant nation-state, rather than with the individual or indigenous community whose land or property houses the relevant genetic resource. Under the Common System, the relevant state either expressly owns or exercises virtually complete control over such resources and only the state is a party to the benefit-sharing agreement with the bioprospector. See in general the WHO: *Genomics and world health: Report of the advisory committee on health research* (2002), available at <http://bit.ly/lpiMYg> (accessed 24 April 2011) 143. For a text of the Andean Pact, see <http://bit.ly/mx5o5L> (accessed 25 April 2011). Other examples are the regional African Model Law for the Protection of the Rights of Local Communities, Farmers and Breeders, and for the Regulation of Access to Biological Resources of 2000, referred to above in fn 33, and the Asean Regional Framework Agreement on Access to Biological and Genetic Resources (2004).

64 Hamilton "The Human Genome Diversity Project and the New Biological Imperialism" 2001 (41) *Santa Clara LR* 619 636–639.

65 See in general WHO *Genomics and world health* fn 63 above.

66 The discussion under para 3 draws to some extent on a discussion of related issues by Slabbert and Pepper 439–444.

organisations, such as the Human Genome Organisation (HUGO),⁶⁷ and finally bodies that represent industrialised nations, such as the OECD.⁶⁸

3.1.1 *Convention on Biological Diversity*⁶⁹

The 1992 UN Convention on Biological Diversity created a regulatory framework for the access and use of genetic resources. Member countries commit to conserve, sustainably develop, and equitably share benefits resulting from biological resources. Article 15 of the Convention on Biological Diversity recognises the sovereign rights of states over their natural resources, including their authority within the respective national jurisdictions, to determine access to genetic resources.⁷⁰ Article 15 also sets out the conditions for access to genetic resources. These involve not only the recognition of state sovereignty, already referred to, but also facilitation of access by other states; prior informed consent of the state of origin of the genetic material and fair sharing of the results of research relating to the use of the genetic resources.

Safrin,⁷¹ however, argues that nations, in the aftermath of the CBD, instead of promulgating laws to enable and facilitate access to genetic resources, have curtailed access by enacting laws to restrict access to their genetic resources. An access-restricting regime will undoubtedly deter research activity in genetic-rich jurisdictions for a range of reasons, which may include reluctance to engage with or violate arduous access-restricting provisions in developing countries.

Article 19 of the CBD refers to the distribution of benefits and enjoins states to provide for effective participation in the research on genetic resources originating from that state.⁷² Although the Convention was adopted in 1992 and entered into force in 1993, the putting into operation of these provisions only began in earnest in 1999, leading to the Bonn Guidelines on Access to Genetic Resources and the Fair and Equitable Sharing of Benefits arising out of their

67 See eg the official website at <http://www.hugo-international.org>.

68 Organisation for Economic Co-operation and Development; see <http://www.oecd.org>.

69 Signed by 150 government leaders at the 1992 Rio Earth Summit, dedicated to promote sustainable development. The Convention, negotiated under the auspices of the United Nations Environment Programme (UNEP), entered into force on 23 December 1993. The three goals of the CBD are to promote the conservation of biodiversity, the sustainable use of its components, and the fair and equitable sharing of benefits arising out of the utilisation of genetic resources.

70 Article 15 reads: "1. Recognising the sovereign rights of States over their natural resources, the authority to determine access to genetic resources rests with the national governments and is subject to national legislation. 2. Each Contracting Party shall endeavour to create conditions to facilitate access to genetic resources for environmentally sound uses by other Contracting Parties and not to impose restrictions that run counter to the objectives of this Convention." It is clear, however, that there are asymmetries regarding how individual countries have implemented the provisions relating to access and benefit-sharing in the CBD and the resulting Bonn Guidelines. See in this regard Ghose *Benefit sharing of genetic resources* (fn 24 above) 7.

71 Safrin 664 fn 212.

72 One problem with this provision is that the Convention fails to address the issue of genetic material held in international biobanks which have been sourced over many years from local communities in developing countries. See Downes "New diplomacy for biodiversity trade: Biodiversity, biotechnology and intellectual property in the Convention on Biological Diversity" in Hunter, Salzman & Zaelke (eds) *International environmental law and policy* (2002) 945.

Utilization,⁷³ aimed at assisting parties, governments and stakeholders in developing access and benefit-sharing strategies when establishing legislative, administrative and policy measures or negotiating contractual agreements relating to access and benefit-sharing.⁷⁴ However, as is the case with the CBD, the Bonn Guidelines are not legally binding unless these have been incorporated into domestic legislation. South Africa ratified the Convention in November 1995 and subsequently enacted the National Environmental Management: Biodiversity Act⁷⁵ in 2004.

The question whether the Convention can be used as a model for the regulation of access and benefit-sharing of human genetic material, raised above, is not a simple one. Firstly, if the CBD is used as a blue-print for the regulation of human genetic resources, definitions should be unambiguous. The term “indigenous” in the CBD, for example, is nowhere defined and also not in the Bonn Guidelines, as well as South Africa’s Biodiversity Act, the latter referring to “indigenous community” in section 82.⁷⁶ The term “traditional knowledge” referred to in the CBD and Bonn Guidelines is another example.⁷⁷

73 Adopted by the parties to the Convention at its 6th meeting held in The Hague in April 2002. For a copy of the Guidelines, see <http://bit.ly/jfi3Ys> (accessed 19 April 2011).

74 The Bonn Guidelines set up a voluntary framework for legislative, administrative or policy measures on access and benefit-sharing as well as access and benefit-sharing contracts and agreements. The Guidelines further harmonised the steps for adequate access and benefit-sharing and clarified and complemented existing obligations under the CBD. The Bonn Guidelines contain five chapters. The General Provisions (I) include objectives, the scope and definitions. The core of the Guidelines is contained in chapter II, which clarifies roles and responsibilities. The participation of stakeholders is detailed in chapter III. Part IV describes the steps of the access and benefit-sharing process. The other provisions (V) are followed by two appendices that suggest elements for material transfer Agreements (Appendix I) and exemplify monetary and non-monetary benefits (Appendix II). The Guidelines assign different responsibilities to the countries of origin of genetic resources, providers, users and countries with users of genetic resources under their jurisdiction. With regard to benefit-sharing, the Bonn Guidelines cautiously state that mutually agreed terms could cover conditions, procedures and types of benefits to be shared, depending on what is regarded as fair and equitable in the given case (45–50 BG). Annex II of the Guidelines provides examples of monetary and non-monetary benefits. These benefits can be long, medium and short-term. A balance should be achieved between the different forms of benefits. Possible benefits are to be distributed between all those that have contributed to the research and resource management, such as government, academic institutions and indigenous communities. The Guidelines also stipulate that benefits should be “directed in such a way as to promote conservation and sustainable use of biological diversity”. See also Dross & Wolff *New elements of the international regime on access and benefit-sharing of genetic resources: The role of certificates of origin* (BfN: Federal Agency for Nature Conservation, 2005), available at <http://bit.ly/IE5yaK> at 17 (accessed 29 April 2011).

75 10 of 2004. The objectives of this Act, stated in s 2 *inter alia* provide for the management and conservation of biodiversity within the Republic; the use of indigenous biological resources in a sustainable manner and the fair and equitable sharing among stakeholders of the benefits that arise from bioprospecting that involves indigenous biological resources.

76 Eg a 8(j) of the CBD.

77 The term does not appear in the Biodiversity Act, which instead refers to “traditional uses”, eg s 82(1)(b).

3 1 1 1 What is an “indigenous community” for the purpose of benefit-sharing agreements?

Finding an acceptable definition for an “indigenous community” is a challenge, as is observed in the 2004 UN document, *The concept of indigenous peoples*, by the UN Department of Economic and Social Affairs.⁷⁸ The prevailing view is that no formal universal definition of the term is possible (or necessary)⁷⁹ and that self-definition by a specific community as an “indigenous community” is an important criterion.⁸⁰ Applied to the South African context, such a determination of what constitutes an indigenous community (in contrast to a non-indigenous one), could be problematic. In South Africa’s legal discourse, the term “indigenous” is used in reference to the languages and legal customs of the majority of the black African population as opposed to the other races. Similar to the view held in the rest of Africa, there is no criterion for identifying indigenous peoples in South Africa.⁸¹ Crawhall⁸² refer to four main groups self-identifying

78 See UN, Division for Social Policy and Development Secretariat of the Permanent Forum on Indigenous Issues, *The concept of indigenous peoples* (Background paper prepared by the Secretariat of the Permanent Forum on Indigenous Issues) PFI/2004/W.S1/3 (New York, January 2004) 4.

79 *The concept of indigenous peoples* 4.

80 One of the most cited descriptions of the concept of the indigenous was given by Jose R Martinez Cobo, the Special Rapporteur of the Sub-Commission on Prevention of Discrimination and Protection of Minorities, in his famous *Study on the problem of discrimination against indigenous populations* (UN Doc E/CN.4/Sub.2/1986/7 and Add. 1–4, paras 379–382), which states as follows: “Indigenous communities, peoples and nations are those which, having a historical continuity with pre-invasion and pre-colonial societies that developed on their territories, consider themselves distinct from other sectors of the societies now prevailing on those territories, or parts of them. They form at present non-dominant sectors of society and are determined to preserve, develop and transmit to future generations their ancestral territories, and their ethnic identity, as the basis of their continued existence as peoples, in accordance with their own cultural patterns, social institutions and legal system.” Furthermore: “This historical continuity may consist of the continuation, for an extended period reaching into the present of one or more of the following factors: (a) Occupation of ancestral lands, or at least of part of them; (b) Common ancestry with the original occupants of these lands; (c) Culture in general, or in specific manifestations (such as religion, living under a tribal system, membership of an indigenous community, dress, means of livelihood, lifestyle, etc.); (d) Language (whether used as the only language, as mother-tongue, as the habitual means of communication at home or in the family, or as the main, preferred, habitual, general or normal language); (e) Residence on certain parts of the country, or in certain regions of the world; (f) Other relevant factors.” An indigenous person is hence one “who belongs to these indigenous populations through self-identification as indigenous (group consciousness) and is recognized and accepted by these populations as one of its members (acceptance by the group) . . . This preserves for these communities the sovereign right and power to decide who belongs to them, without external interference”.

81 See Crawhall *Indigenous peoples of South Africa: Current trends* (Project to Promote ILO Policy on Indigenous and Tribal Peoples/South African San Institute, International Labour Office, Geneva, 1999) 2–11; Chennels & Du Toit “The rights of indigenous peoples in South Africa” in Hitchcock and Vinding (eds) *Indigenous peoples’ rights in Southern Africa* IWGIA (2004) 98. See in particular Mukundi *South Africa: Constitutional, legislative and administrative provisions concerning indigenous peoples* Centre for Human Rights, University of Pretoria (Country report of the research project by the ILO and the African Commission on Human and Peoples’ Rights on the constitutional and legislative protection of the rights of indigenous peoples: South Africa) (ILO 2009).

82 Crawhall 5–7.

as indigenous peoples in South Africa: the San who comprise various sub-ethnic groups (!Xû and Khwe, ‡Khomani, /'Auni, Saasi, //Xegwi, !Kung, and /Xam descendants); the Nama who are also referred to as the Khoekhoean; Griqua and Korana; and revivalist Khoesan groups.

Schuklenk and Kleinsmidt⁸³ argue that even if an indigenous community can be delineated for the purpose of benefit-sharing, the problem of representation of that community remains. For example, the traditional leadership of such community may be inherently undemocratic or exclude women from decision-making as to the types and distribution of whatever benefits may accrue, contrary to *inter alia* the fundamental right to equality and non-discrimination embodied in the South African Constitution.⁸⁴ Researchers may be unclear whether they should maintain neutrality or insist on including their notion of fairness into the negotiations relating to benefit-sharing.⁸⁵

Alternatively, researchers may be confronted with an indigenous community without a recognisable traditional leadership system, as may apply in respect of the San and Khoe leadership, which is yet to be resolved.⁸⁶ Foreign researchers should also be cognisant of the protection of the fundamental rights of indigenous peoples in South Africa, as well as the specific vulnerabilities associated with these groups.⁸⁷

Finally, it is interesting to note that the HUGO Ethics Committee's Statement on Benefit Sharing refers to a "community" in the widest sense,⁸⁸ noting specifically that even "[p]ersons with the same disease could form a community of origin if there is a family history, as may be the case for monogenic disorders (single gene), or a community of circumstance, which is usually the case for common multifactorial diseases."

Another problem with using the CBD as a model for the regulation of access and benefit-sharing of human genetic resources is that of genomic sovereignty of human genetic resources. As was alluded to above,⁸⁹ issues regarding the control and ownership of human genetic material are more controversial than that of genetic material in agricultural resources and should be very carefully considered.

83 Schuklenk & Kleinsmidt "North-south benefit sharing arrangements in bioprospecting and genetic research: A critical ethical and legal analysis" 2006 (6)3 *Developing World Bioethics* 122126–127.

84 Of the Republic of South Africa, 1996.

85 Schuklenk & Kleinsmidt 126. At present most indigenous women do not take part in designing and participating in negotiations and strategies to uplift their communities. Indeed, as Crawhall 33 remarks: "[M]ost spokespersons recognised by the government tend to be men and in effect male agendas dominate planning and negotiations."

86 See eg Mukundi 12 who states that the structures of these communities had been completely destroyed so that they are not recognised by the new government. There are, however, efforts to recognise and re-establish the leadership structures in these communities.

87 Mukundi 21–62. These relate to recognition and identification as an indigenous community; non-discrimination; participation in decision-making and consultation; self-management; access to justice; cultural and language rights; education; land, natural resources and environment; socio-economic rights, etc. See *inter alia* ss 9, 15, 30 and 31 of the Constitution.

88 See fn 21 above.

89 See para 2 2 2 1 above.

3 1 2 *The Universal Declaration on the Human Genome and Human Rights; the International Declaration on Human Genetic Data and the Universal Declaration on Bioethics and Human Rights*

The Convention on Biological Diversity referred to above,⁹⁰ although recognising the territorial sovereignty of states over genetic resources within their jurisdictions, does not explicitly extend this to genetic resources of *human* origin, as explained above.

UNESCO's Universal Declaration on the Human Genome and Human Rights⁹¹ is prospective in nature and provides the basic ethical principles for the proper conduct of human genome research generally. The Declaration emphasises the inherent dignity of all persons, regardless of their genetic characteristics,⁹² as well as the idea that the human genome – in its natural state – should not give rise to financial gain.⁹³ Benefits gained from research on the human genome should be made available to all (the notion of “access”), and the applications of this research should seek to offer relief from suffering and aim to improve the health of individuals and humankind in general.⁹⁴ Although not legally binding, the Declaration represents the dynamic development of international legal norms and reflects the commitment of member states to abide by certain principles. As a form of “soft law”, these principles may in time become entrenched as exacting standards, as there is no doubt that the Declaration has already significantly affected the *opinio iuris* of the international community.⁹⁵

In 2003, UNESCO adopted the International Declaration on Human Genetic Data⁹⁶ which focuses on the protection of human genetic data under international human rights law. The aims of the Declaration are *inter alia* to ensure that the collection, processing, use and storage of human genetic data, human proteomic data and biological samples conform with requirements for respect for human dignity and the protection of human rights and fundamental freedoms.⁹⁷

UNESCO's Universal Declaration on Bioethics and Human Rights, adopted in 2005,⁹⁸ reiterates similar principles and specifically aims to promote equitable access to medical, scientific and technological developments. It also emphasises the importance of the free flow and rapid sharing of knowledge concerning these developments, as well as the sharing of benefits, with particular emphasis on the needs of developing countries.⁹⁹ In the case of transnational research, the

90 See para 2 above.

91 Adopted 11 November, 1997, text available at <http://bit.ly/jueX7K> (accessed 5 March 2010).

92 A 2 of the Declaration.

93 A 4.

94 A 12.

95 See Francioni *Genetic resources, biotechnology and human rights: The international legal framework* (European University Institute Working Papers, EUI LAW No 2006/17) 8, available at <http://bit.ly/182auv> (accessed 5 March 2010).

96 Adopted 16 October, 2003.

97 A 1.

98 Adopted 19 October, 2005, text available at <http://bit.ly/k57fOG> (accessed 5 March 2010).

99 A 2(f). See also a 15. See in this regard Hardy *et al* “The next steps for genomic medicine: Challenges and opportunities for the developing world” October 2008 *Nature Reviews Genetics* S23–S27.

Declaration advises that this research should be responsive to the needs of host countries and to the alleviation of urgent global health problems.¹⁰⁰

3 1 3 *World Health Organisation: Indigenous Peoples and Participatory Health Research (2002)*

The World Health Organisation's document on Indigenous Peoples and Participatory Health Research¹⁰¹ provides useful instruction regarding the joint management of research by research institutions and indigenous peoples. The document captures the most significant provisions of an ideal agreement between research institutions and indigenous peoples, drawing on experiences in various countries and providing references to key literature.

Although not legally binding, this document testifies to some of the problems relating to genomic sovereignty, particularly in instances where genomic research is undertaken with DNA samples of local indigenous populations, without any regard to the balancing of the interests, benefits and responsibilities between the researchers and the indigenous peoples involved in the research. Mechanisms for ethical review may be weak or non-existent in some developing countries, whereas low levels of education and cultural or language barriers may increase the likelihood of exploitation of indigenous peoples and their genetic material.

3 1 4 *The Recommendation on Research on Biological Materials of Human Origin of 2006*

The Council of Europe's Recommendation on Research on Biological Materials of Human Origin,¹⁰² the first instrument addressing genomic databases at a supranational level, provides in Chapter V for the regulation of and access to "population biobanks",¹⁰³ a term that denotes a large repository or collection of human DNA from which genomic databases will be derived.¹⁰⁴ It is worth noting that countries that have or are in the process of establishing large genomic databases are Iceland (eg the Icelandic Health Sector Database); Estonia (eg the

100 A 21(3).

101 Accessible at <http://bit.ly/moE5tF> (accessed 5 March 2010).

102 Recommendation Rec (2006)4 of the Committee of Ministers to member states on research on biological materials of human origin (adopted by the Committee of Ministers on 15 March 2006), text available at <https://wcd.coe.int/ViewDoc.jsp?id=977859> (accessed 5 March 2010). Earlier in 1997, the Council of Europe adopted the Convention on Human Rights and Biomedicine (Oviedo, 4 Paris 1997, ETS 164) which, together with its Additional Protocol, addresses broadly the human rights implications of the applications of biology and medicine. The binding force of this Convention is limited to signatory states who have introduced national legislation aligned to the principles of the Convention. The aim of the Convention is to protect the dignity and identity of all human beings and to guarantee everyone, without discrimination, respect for their integrity and other rights and fundamental freedoms with regard to the application of biology and medicine (see A 1).

103 A 17 defines a "population biobank" as "a collection of biological materials that has the following characteristics: (i) the collection has a population basis; (ii) it is established, or has been converted, to supply biological materials or data derived therefrom for multiple future research projects; (iii) it contains biological materials and associated personal data, which may include or be linked to genealogical, medical and lifestyle data and which may be regularly updated; and (iv) it receives and supplies materials in an organised manner."

104 See Lowrance "The promise of human genetic databases" 2001 *BMJ* 1009.

Estonian Genome Project Gene Bank); the United Kingdom (UK Biobank); the USA (National Cancer Institute Biorepository); Scotland (the Generation Scotland projects) and Canada (the CARTaGENE Project). Databases of this nature, if established soon for African populations in Southern Africa or South Africa specifically, could ultimately serve the whole continent.

The Recommendation instructs that a proposal to establish a population biobank should be subject to an independent examination of its compliance with the provisions of this recommendation;¹⁰⁵ and that each population biobank should be subject to independent oversight in order to safeguard the interests and rights of the persons concerned in the context of the activities of the biobank.¹⁰⁶ Despite the European focus of this Recommendation, it nevertheless provides useful guidelines to South Africa regarding the setting up of a population biobank.

3.1.5 HUGO's Statement on Human Genomic Databases of 2002

The Human Genome Organization's Statement on Human Genomic Databases of 2002¹⁰⁷ recommends *inter alia* that: (1) Human genomic databases are global public goods; (2) Knowledge useful to human health belongs to humanity; (3) Human genomic databases are a public resource; and that (4) All humans should share in and have access to the benefits of databases.¹⁰⁸ Despite the fact that this Statement is not legally binding on countries, HUGO has significant influence over the international scientific community.¹⁰⁹

Knoppers, Abdul-Rahman and Bédard¹¹⁰ rightly note that the description of a population biobank contained in the Council of Europe's Recommendation on Research on Biological Materials of Human Origin differs from that defined in HUGO's Statement on Human Genomic Databases of 2002.¹¹¹ Terms such as "biobank"¹¹², "gene bank"¹¹³, "tissue bank" and "human genetic research databases" are interchangeably used in the context of population biobanks. It is also not clear whether these banks may contain samples alone, information alone, or linked combinations of the two.

The lack of universally agreed upon terminology relating to population databases illustrates that there is still much confusion over the very nature of, risks and benefits relating to population (genomic) research, the latter being very

¹⁰⁵ A 18.

¹⁰⁶ A 19.

¹⁰⁷ Adopted December 2002, text available at <http://www.eubios.info/HUGOHGD.htm> (accessed 5 March 2010). The Human Genome Organisation (HUGO) was established in 1989 as an international organisation, primarily to foster collaboration between scientists around the world that are working on the human genome.

¹⁰⁸ Rec 1.

¹⁰⁹ Recs 1 and 2.

¹¹⁰ "Genomic databases and international collaboration" 296.

¹¹¹ Which defines a "genomic database" as being simply "a collection of data arranged in a systematic way so as to be searchable". See HUGO Ethics Committee, Statement on Human Genomic Databases.

¹¹² Eg used in the United Kingdom, eg the Biobank UK project, at www.ukbiobank.ac.uk, and in Sweden, the Swedish National Board of Health and Welfare biobanks in medical care, at <http://bit.ly/k0MTeS> (accessed 6 March 2010).

¹¹³ As established via the Estonia Genome Project, see Human Genes Research Act 2000 (Estonia) section 2(10).

different from those related to other databases (eg of residual tissue collections or biosamples collected during clinical trials).¹¹⁴ This position is compounded by a range of different bodies issuing partly overlapping documents or guidelines in an uncoordinated manner.¹¹⁵ Despite the fact that these databases may differ in their intended uses, the unifying element is that they have been primarily created for the purposes of medical or other human research, unlike that of archived pathology samples, for example.

3.1.6 *OECD and others*

The Organisation for Economic Co-operation and Development (OECD), established in December 1960 by virtue of the Convention on the Organisation of Economic Co-operation and Development, has as one of its primary aims the promotion of policies designed “to achieve the highest sustainable economic growth and employment and a rising standard of living in Member countries”. In 2009, the OECD Council issued the OECD Guidelines for Human Biobanks and Genetic Research Databases (HBGRDs),¹¹⁶ provides valuable guidelines relating to the establishment, management, governance, operation, access to, use and discontinuation of HBGRDs. It recognises that one of the fundamental objectives of an HBGRD is to foster scientific research. It specifically seeks to facilitate wide access to data and materials for biomedical research, as well as to ensure that research is conducted in a manner respectful of participants, and that it upholds human dignity, fundamental freedoms and human rights. The Guidelines make specific and culturally sensitive recommendations regarding custodianship of HBGRDs and benefit-sharing.¹¹⁷ South Africa is not a member of the OECD but is classified as an “Enhanced Engagement Country”.

As already mentioned above, despite the fact that the OECD’s guidance documents lack enforceability, as membership is limited, these guidelines, via the OECD as a forum for reaching multilateral policy agreement, could assume potentially powerful, global norm-setting standards.

The Canadian Institutes of Health Research (CIHR) issued Guidelines for Health Research involving Aboriginal People in 2007.¹¹⁸ These Guidelines recommend the co-ownership of samples and data between researchers and Aboriginal people involved in research;¹¹⁹ and that biological samples used by researchers should be considered “on loan” to the researchers unless otherwise specified in the research agreement, eg a researcher is considered to be a steward

114 Samples may be found in “private collections”, which are procured directly or indirectly from a donor by a researcher, who uses the samples immediately or store them for later use; “study-associated collections” which are collected in the course of a particular research project, study or trial; diagnostic pathology department archives where sample collections exist as a result of normal health care practices; sample resources arising as a by-product of transplantation or transfusion; and sample collections specifically set up to service as a human biological bank.

115 Other related HUGO Statements are, for example, the Ethical, Legal, and Social Issues Committee Statement on the Principled Conduct of Genetic Research of 21 March, 1996; and the HUGO Ethics Committee Statement on Benefit Sharing of 9 April 2000.

116 Accessible at <http://bit.ly/jn4Tru> (accessed 6 March 2010).

117 See Rec 9.

118 Accessible at <http://bit.ly/jDnuC5> (accessed 6 March 2010).

119 A 12.5.

rather than as owner of the relevant samples.¹²⁰ The Guidelines recognise the importance of special protection for indigenous cultural (and also sacred) knowledge, but do not refer specifically to genomic sovereignty or ownership of genomic information identified for specific studies, apart from the recommendation regarding co-ownership of samples between participants and researchers.¹²¹

Another instructive document that advises on genomic research, access and benefit-sharing is the Human Genome Diversity Project's Model Ethical Protocol for Collecting DNA-samples of 1997,¹²² which provides detailed guidance on how to approach and plan genomic research, eg by learning as much as possible about the cultural, social and religious practices of the relevant population.¹²³ The Protocol states that the Project will not profit from any commercial uses of samples it gathers or knowledge derived from these samples, and that should commercial products be developed as a result of the Project's collections, a "fair share" of the financial rewards must return to the sampled populations. Researchers who take part in the Project must accept these two points.¹²⁴

3.1.7 Preliminary conclusion

The brief overview of the various statements, guidelines and documents above shows that these documents address more general principles and activities and none have a clear mandate or authority to formulate an internationally accepted position or norms and standards to oversee the governance of international collaborative genomic research.¹²⁵ The status, authority, content and enforceability mechanisms of these bodies differ widely. The general approach is a cautious one; the result of a public perception that researchers are "playing God" or tampering with God's creation; of visions of clone farms, organ banks and the creation of robot-like beings.¹²⁶ The notions of access to and benefit-sharing of genetic resources are generally recognised in most of these documents in some or other format.

120 A 13.

121 Aa 7 and 8. A 8 states that any intellectual property claims should be addressed in the negotiations prior to the commencement of the research project.

122 North American Regional Committee. Text available at <http://bit.ly/IRuNdx> (accessed 7 March 2010). The Human Genome Diversity Project is an international effort to collect, preserve, analyse, and make available genetic and ethnographic information from people around the world.

123 In some societies, hair is secretly collected from intended victims to harm them through witchcraft. Consequently, these people may collect their own loose hair, fingernail parings, and other body products and bury them to avoid this danger. Researchers who ask such a population for hair may be perceived as intending to perform witchcraft. Blood is often intended to be used as a sacrifice, sometimes through special rituals. Donation of blood in some cultures is a serious issue that may require discussion and perhaps a neutralising ritual. Before approaching the population, researchers need to know as much as possible about its likely concerns about and reaction to their plans for sample collection. See para II of the Protocol.

124 Para IX of the Protocol.

125 Knoppers, Abdul-Raham & Bedard 294.

126 Moore "Owning genetic information and gene enhancement techniques: Why privacy and property rights may undermine social control of the human genome" 2000 14(2) *Bioethics* 97-98.

3 2 South African context

Access to genetic resources in terms of the CBD is determined by the various national member governments and is subject to national legislation,¹²⁷ which, together with the access to, and sharing of, the benefits of genetic information are the two cross-cutting themes of the CBD. South Africa's Biodiversity Act of 2004 also regulates activities described as "bioprospecting",¹²⁸ including access to and benefit-sharing of indigenous biological resources. The definitions of "genetic material" and "genetic resources" in both the CBD and the Biodiversity Act, as well chapter 6 and section 80(2)(b)(i) of the Biodiversity Act clearly exclude genetic material of *human* origin.¹²⁹ As referred to above, although access to and benefit-sharing of human genetic material are excluded by the Biodiversity Act, the Act's regulation of this in terms of indigenous biological resources may be instructive as regards future formulation of the legal regulation of access to and benefit-sharing relating to human genetic material.¹³⁰

Chapter 6 of the Act provides that no bioprospecting may be carried out and no biological material be exported from the Republic without a relevant permit.¹³¹ Section 82 provides for the protection of interests of potential stakeholders in a bioprospecting project, as well as for the conditions for the issuing of such a permit, which include the requirement that the applicant concludes a material transfer agreement (MTA) regulating provision of access to the biological resource, as well as a benefit-sharing agreement in terms of which the "stakeholder" (which could be a person, organ of state, community or indigenous community providing access to the biological resource) shares in future benefits arising from the bioprospecting.¹³² The Minister¹³³ must also approve such benefit-sharing agreement.¹³⁴ If the stakeholder is an indigenous community, similar requirements apply.¹³⁵ Chapter 6 furthermore provides that the issuing authority may facilitate negotiations between the applicant and stakeholder and ensure that this is conducted in a fair and equitable manner and on equal footing.¹³⁶

Section 83 stipulates the legal requirements regarding a benefit-sharing agreement. These requirements, *inter alia*, state that the manner in which, and

127 The CBD defines in a 2 ("use of terms") the phrase "genetic resources" as "genetic material of actual or potential value", whereas "genetic material" which is used in the latter definition is described as "any material of plant, animal, microbial or other origin containing functional units of heredity".

128 "Bioprospecting" is defined as "research on, or development or application of, indigenous biological resources for commercial or industrial exploitation".

129 S 80(2)(b)(i) of the Biodiversity Act states that "indigenous biological resources" excludes "genetic material of human origin". The Chapter referred to is Chapter 6 (Bioprospecting, access and benefit sharing). This also applies in respect of the Bonn Guidelines and Nagoya Protocol, referred to above.

130 Benefit-sharing, for example, is mentioned by H3Africa (Human Heredity and Health in Africa) in its White Paper: *Harnessing genomic technologies toward improving health in Africa: Opportunities and challenges* (January 2011) 30–31.

131 S 81.

132 S 82(2).

133 Responsible for national environmental management.

134 S 82(2)(c).

135 S 82(3).

136 S 82(4).

the extent to which the relevant indigenous biological resources are to be utilised or exploited for the intended bioprospecting, as well as the manner in which and the extent to which, the stakeholder will share in the benefits that may arise from such bioprospecting be specified in such an agreement.¹³⁷ No benefit-sharing agreement becomes effective until approved by the Minister.¹³⁸ Chapter 6 also regulates the conclusion of material transfer agreements.¹³⁹

4 BENEFIT-SHARING IN PRACTICE

The adoption of benefit-sharing may take any of three forms: corporate, state or contractual.¹⁴⁰ The advantage of a contractual approach in respect of benefit-sharing is that it is respectful of the autonomy of donors by providing input and negotiating power in the context of commercial genetic research, especially where an individual's biological material may have commercial value or where there is a homogenous disease community.¹⁴¹ It will be less useful for large-scale genetic databases where the donor group is more heterogeneous. Private contractual agreements may require state intervention where an interpretation of "freedom of contract" has the potential to perpetuate inequality, exploitation and power imbalances.¹⁴² The contract model, known as the Biotrust model and developed by Winickoff and Neumann, attempts to promote community involvement and participation in genomic governance. Following on the idea of the Charitable Trust Model (first proposed by Winickoff and Winickoff in 2003), the central idea of the Biotrust model is to use the charitable trust as the legal framework for genomic biobanks.¹⁴³ In this model, donors transfer their property interests, if relevant, through a series of trust instruments to the same trustee (the Biotrust Foundation) which would hold title to the trust property for the benefit of the beneficiary. The public, in the case of the charitable trust, is the beneficiary,¹⁴⁴ who would assist with defining the charitable purpose and the terms of public benefit. The Foundation would be managed by its own by-laws, to which the donors must agree, which would specify that the use of the trust property will be contingent on review approval of two bodies, namely the Ethics Review Committee and the Donor Advisory Committee.

The organising principle of the trust would be that of public benefit and not profit. In terms of the trust model, the trustees would negotiate shared intellectual property arrangements with researchers with the purpose of using a portion of profits so derived to finance the operation of the trust.¹⁴⁵ The trustees could evaluate research applications according to criteria that ensure public

137 S 83(1).

138 S 83(2).

139 S 84.

140 Knoppers (2000) at 214.

141 Hunter "Benefit-sharing: From principle to pragmatism addressing public concerns about commercialization in population genetic research", Arts and Humanities Research Council Centre for Studies in Intellectual Property and Technology Law, available at <http://bit.ly/md04co>, 1753 1755.

142 *Idem* 1756.

143 Winickoff & Neumann "Towards a social contract for genomics: Property and the public in the 'Biotrust' model" 2005 (1)3 *Genomics, society and policy* 8 10.

144 Hunter 1756.

145 *Ibid.*

benefit, and the allocation of benefits could be directed towards either the general public or specific groups.

One advantage of the trust model is that it can be applied on a project, institutional, national or regional basis.¹⁴⁶ Initial funding for the infrastructure of such a trust could be supplied by institutions with a stake in public health, eg health departments, national research funding bodies, and foundations. Access fees for tissues requiring exportation, for example, could be levied. The processes involved in the creating, maintaining and providing biological samples are costly and the levying of fees in the provision of the samples would be an acceptable and legitimate practice.

It is not inconceivable that the volume of tissues that could be exported from southern Africa may be substantial, if the exportation of human biological samples from Uganda is considered by way of comparison.¹⁴⁷ Partnership with the private sector would also be a possible funding avenue. As an investment tool to build up local capacity, the trust model has many advantages and is indeed a viable model in the context of the human genetic research in South Africa. The Biodiversity Act of 2004 already refers to the state's "trusteeship" of biological diversity.¹⁴⁸ It is also recommended that in South Africa, samples and information be held in a form of "trusteeship", or alternatively, custodianship or stewardship.

One practical problem with this model, however, is that it may be difficult to implement where the volunteer or donor group is more heterogeneous and disparate. Most of the potential problems could be circumvented by careful drafting of the relevant by-laws before the framework of the model is set up, ideally through a process of public consultation.¹⁴⁹ This would provide a means for public input into the process of benefit-sharing.

It is worth noting that as an alternative to these benefit-sharing models, a tissue tax model has been suggested. This tax is based on the assumption that the human genome (and indirectly also human tissue and cells) constitute a national resource. The author of this idea, Bovenberg,¹⁵⁰ suggests that the tax be levied on profits made on a gene patent, for example, or on a biological product developed from human sources. The advantage of this tax model is that it is enforceable and collectible and that it applies regardless of whether persons have specific genes or whether they participated in the research; it preserves existing incentives, such as patents; avoids speculation over the relative contributions of specific tissue to the end-product; is only due when actual profit is made; and its collection and redistribution is subject to democratic control. It would also prevent complicated, protracted negotiations and would allow the larger community to share in the benefits that may arise.¹⁵¹ The almost immediate response to this "DNA-tax" is that it would treat human DNA as a taxable commodity. Hunter¹⁵² lists a number of justifiable drawbacks of this model, of which one is that such system would

146 Emerson *et al* "Access and use of human tissues from the developing world" 3.

147 Upshur, Lavery & Dindana "Taking tissue seriously means taking communities seriously" 2007 (8) *BMC Medical Ethics* 11.

148 S 3.

149 Winickoff & Neumann 15.

150 Bovenberg *Property rights in blood, genes and data: Naturally yours?* (2006) 200.

151 Hunter 1757.

152 *Ibid.*

not accommodate access to the spin-offs of biomedical research, such as new drugs or therapies. Another disadvantage of this model is that it does not involve the public or community in deciding on how the benefits are to be distributed.

4 1 Key elements in benefit-sharing agreements

The following elements may be considered as key elements in benefit-sharing agreements:¹⁵³

4 1 1 *Prior to the conclusion of the benefit-sharing agreement*

Issues such as relationship building, a negotiation framework, as well as capacity funding for the negotiations are relevant. Face-to-face interactions will assist relationship building, as well as establishing a negotiation framework where a draft term sheet or memorandum of understanding could be considered. It is during this stage that finality should be reached on whether the benefit-sharing agreement should be legally binding or non-binding.

4 1 2 *Determining the contract parties*

The determination of the contact parties will depend on whether the project may affect rights or title. The agreement should be concluded with the official representatives (whether elected or hereditary); in the case of government parties, with those with the responsibility to sign the contract off. In the case of corporate or other parties, the appropriate legal entity or entities (eg corporation, subsidiary, non-profit society) should be determined, to ensure authorised signatories. Provisions should be included in the agreement to bind successor companies.

4 1 3 *Background and foundational principles*

The agreement could contain appropriate “whereas” clauses, which in the case of developing countries as one of the contract parties, may include reference to capacity building and resource allocation. The mandate and objectives of both contract parties should be stated.

4 1 4 *Definitions*

For the sake of transparency and legal clarity, all terms related to benefit-sharing and access should be clearly defined.

4 1 5 *Agreement purpose*

There should be certainty in respect of exactly what the agreement will provide and what each party will provide and receive in return. Non-derogation of specific rights, if relevant, could be included. This component, specifically, could include the following:¹⁵⁴

- Description of genetic resources covered by the material transfer agreements, including accompanying information;
- permitted uses, bearing in mind the potential uses of the genetic resources, their products or derivatives under the material transfer agreement;

¹⁵³ Guidance in this regard is based on the guide for *Benefit-sharing agreements in British Columbia: A guide for first nations, businesses, and governments* (2009), available at <http://bit.ly/iMp3za> (accessed 29 April 2011).

¹⁵⁴ See *Bonn Guidelines* (Appendix I).

- a statement that any change of use would require new prior informed consent and material transfer agreement;
- whether intellectual property rights may be sought and if so, under which conditions;
- terms of benefit-sharing arrangements, including commitment to share monetary and non-monetary benefits;
- no warranties guaranteed by provider on identity and/or quality of the provided material;
- whether the genetic resources and/or accompanying information may be transferred to third parties and if so, conditions that should apply; and
- the duty to minimise environmental impacts of collecting activities.

Specific legal provisions also relevant under this heading, as prescribed by the *Bonn Guidelines*, are:

- The obligation to comply with the material transfer agreement;
- the duration of agreement;
- notice to terminate the agreement;
- that the obligations in certain clauses survive the termination of the agreement;
- the independent enforceability of individual clauses in the agreement;
- events limiting the liability of either party (such as act of God, fire, flood, etc);
- dispute settlement arrangements;
- assignment or transfer of rights;
- assignment, transfer or exclusion of the right to claim any property rights, including intellectual property rights, over the genetic resources received through the material transfer agreement;
- a choice of law-clause;
- a confidentiality clause; and a
- guarantee.

4 1 6 Administration and implementation costs

This section sets out what each party needs in order to implement the agreement (eg management support; employment liaison; accounting support, etc).

4 1 7 Communications

This part details the contact persons for each party, as well as the process for ensuring timeous communication throughout the project. Information requirements should also be specified.

4 1 8 Decision-making

The exact nature of the decision-making process should be detailed, eg will there be an advisory role or consultation; will it be full-shared decision-making, or something in between? The type of decisions, and the type of processes by which they are to be reached, should be explained, distinguishing between major decisions, operating decisions, business decisions, etc.

4 1 9 Possible business of employment benefits

Apart from revenue-sharing, business opportunities may arise (eg for equity, joint ventures; service and supply contracts). Employment opportunities may be relevant if the infrastructure accommodates this. Issues such as training, dispute settlement, etcetera, should be considered.

4 1 10 Community benefits and resources

This category may include a wide range of options that the parties may agree to, such as developments regarding infrastructure; recreation; school or health facilities or supplies; contributions to community events; scholarships, and others.

4 1 11 Terms of the benefit-sharing agreement

The start and end date and relevant termination process should be clearly explained. Any periodic reviews and a process for necessary amendments should also be detailed. It is important to include provisions regarding enforcement of the contract and dispute resolution. The terms should ideally also refer to issues of confidentiality and be cognisant of the relevant state's freedom of information legislation; corporate shareholder disclosure regulations, and the protection of indigenous rights (cultural and otherwise). Standard contract clauses should also be incorporated.

4 1 12 Signing authority and ratification

The authority of the representatives to sign the contract should be verified.

5 CONCLUSION

There is presently no comprehensive and consistent legal network at an international level that governs, in a satisfactory manner, issues relating to genomic sovereignty, genomic databases, access and benefit sharing in respect of human genetic material. There is also no global consensus on the issues of ownership and commercialisation, let alone consistent and uniform definitions relating to access and benefit-sharing.¹⁵⁵ What does emerge, however, is an awareness of (1) the importance of the notions of access and benefit sharing with regard to communities, especially those in developing countries, and (2) the relevant human rights of individual participants, as well as the legal requirements relating to consent.

Genomic studies involving thousands more individuals from Southern Africa¹⁵⁶ are underway, specifically targeting southern Africa's indigenous populations, for example the Zulu, Xhosa, Herero, San and the Sotho-Tswana. In the absence of a clear regulatory framework and no specific insistence on genomic sovereignty regarding human genetic material, including access and benefit sharing in respect of these, there is nothing to protect the genomic information of the South African population from exploitation beyond our borders.

155 Dross & Wolff *New elements of the international regime on access and benefit-sharing of genetic resources: The role of certificates of origin* (BfN: Federal Agency for Nature Conservation, 2005), available at <http://bit.ly/1E5yaK> (accessed 29 April 2011).

156 See Anderson "Kalahari bushmen genome project underway" (26 October 2009) *GenomeWeb Daily News*, available at <http://www.genomeweb.com/print/926414>.

As stated before, whether the Biodiversity Act should be extended to provide for human genomic research and the regulation of access and benefit-sharing of human genetic material as a separate category, or whether separate legislation (with a Human Genetics “Authority” and a special Human Genetic Research Ethics Review Committee) overseeing this should be instituted, is a question that needs to be debated as a matter of urgency.¹⁵⁷ It is clear that the injustice associated with a traditional institutional model that promotes on the one hand participant altruism relating to donations, yet at the same time, sanctions third party commercialisation and private property rights in human genetic material, is untenable.¹⁵⁸ The challenge for South Africa is to find a benefit-sharing model that tempers (not diminishes) commercial interests; that redresses economic imbalance; and that gives research participants a more fair and active role in influencing the sharing of benefits; which will in the long run restore the public trust and confidence in human genetic research.¹⁵⁹

157 Slabbert & Pepper 450.

158 Haddow *et al* “Tackling community concerns about commercialization and genetic research: A modest interdisciplinary model” 2007 (64) *Social Science And Medicine* 272 281.

159 See eg Hunter “DNA as taxable property – the elephant in the room or the red herring?” 2006 (13) *European J of Health Law* 263 264.

Remnants of blameworthiness in the *actio de pauperie*

JC Knobel

BLC LLD

Professor of Private Law, University of South Africa

OPSOMMING

Oorblyfsels van verwytbaarheid in die *actio de pauperie*

Die *actio de pauperie* bevat oorblyfsels van verwytbaarheid van diere. Dit blyk uit 'n analise van die vereistes dat die dier se gedrag *contra naturam sui generis* en *sponte feritate commota* moet wees, asook uit die meerderheid verwere wat teen die aksie opgewerp kan word. Die verwytbaarheid van diere is onvanpas in 'n moderne regstelsel, en *dicta* wat vatbaar is vir 'n vertolking dat diere verwytbaar kan wees, moet vermy word. Risiko-skepping bied 'n aanvaarbare regverdiging vir skuldlose aanspreeklikheid vir skade aangerig deur diere. Die *contra naturam*-vereiste moet afgeskaf word, of konsekwent vertolk word as 'n objektiewe gedragskode wat van regsweë op diere van toepassing gemaak word. Die kousaliteitselement bied 'n sleutel tot die rasonale en regverdige toepassing van die *actio de pauperie*. In die besonder kan juridiese kousaliteit 'n sleutelrol vertolk omdat dit gedrag met verhoogde juridiese relevansie kan beklee sonder verwysing na verwytbaarheid. Die *sponte feritate*-vereiste en verwere teen die *actio de pauperie* moet vertolk word om op kousaliteit betrekking te hê in plaas van verwytbaarheid. 'n Regsvergelykende oorsig oor die posisie in 'n aantal ontwikkelde regstelsels bied steun vir gevolgtrekkings en voorstelle wat in hierdie bydrae gemaak word.

1 INTRODUCTION

The *actio de pauperie* is an action with which the owner of a domestic animal can be held delictually liable for damage caused by the animal. The concept of a “domestic animal” is a fairly wide one which includes, in addition to dogs as may have been expected, such animals as horses, cattle and even bees. The most significant aspect of the action is that it gives rise to strict liability. For this reason, and although the *actio de pauperie* may concur with the *actio legis Aquiliae* and the action for pain and suffering, this action is the most popular remedy in South Africa when harm was caused by an animal belonging to the inclusive category of domestic animals.¹

Because this action is not based on fault, no legal blameworthiness is required on the part of the defendant. The defendant's liability is based on a combination of ownership and the causing of damage by the animal. In a primitive legal society, a partial rationale, if not a comprehensive one, for the *actio de pauperie*

¹ Cf in general Neethling and Potgieter *Neethling-Potgieter-Visser Law of delict* (2010) 357–360; Van der Merwe *Skuldlose aanspreeklikheid vir skade veroorsaak deur diere* (1970) 1 *et seq*; Van der Walt and Midgley *Principles of delict* (2005) 40–41 and case law cited.

was blameworthiness of the animal.² One of the ways in which the owner of the animal could relieve himself of liability was the practice of *noxae deditio*, that is, handing the animal over to the prejudiced person. The latter could then take his wrath out on the animal. The *noxae deditio* fell into desuetude; today the plaintiff's reward is compensation in money.³

One would expect that in the modern legal system of South Africa no blameworthiness is attached to an animal, because an animal cannot understand the law, has no *locus standi*, and is, in fact, not a legal subject in terms of our law. The disappearance of the *noxae deditio*, taken at face value, is a vindication of the idea that the animal cannot be blameworthy because it eliminates the possibility of revenge being taken on the animal, and places the burden to compensate the plaintiff's damage squarely in the hands of a person, the owner of the animal. However, as one takes a more careful look it becomes apparent that, contrary to what one may have expected, remnants of blameworthiness appear to remain embedded in the positive law regulating the application of the *actio de pauperie*. This contribution reflects critically on this phenomenon.

2 REMNANTS OF BLAMEWORTHINESS IN THE SUBSTANTIVE REQUIREMENTS OF THE *ACTIO DE PAUPERIE*

The substantive requirements for successful reliance on the *actio de pauperie* are:

- (a) the animal that causes the damage must be a domestic animal;
- (b) the defendant must be the owner of the animal at the time the damage is caused;
- (c) the animal must behave *contra naturam sui generis* when causing the damage; and
- (d) the plaintiff or his or her property must be lawfully present at the place where and when the damage is caused.

These requirements are generally accepted and can be regarded as trite.⁴ The first two requirements do not need any further attention for the present purpose. However, it is necessary to give consideration to the third and fourth requirements. The third requirement will be reflected upon in this paragraph, and the fourth requirement in the next paragraph dealing with defences to the *actio de pauperie*.

The requirement that the animal must behave *contra naturam sui generis* literally means that it must behave contrary to the nature of its kind when causing damage. Associated with this requirement is a requirement that the animal must behave *sponte feritate commota*, that is, spontaneously and not due to an external influence. Whether or not the owner of the animal can sustain liability for the damage caused by the animal, is thus made to turn on whether the animal behaved *contra naturam sui generis* and *sponte feritate commota* at the time.⁵

² Van der Merwe 1 12 *et seq.*

³ Neethling and Potgieter 357 fn 24; Van der Merwe 1 *et seq.*

⁴ Neethling and Potgieter 358–359; cf Van der Merwe 65 *et seq.*; Van der Walt and Midgley 41 and case law cited.

⁵ Cf in general Neethling and Potgieter 358–359; Van der Merwe 89–93; Van der Walt and Midgley 41 and case law cited.

Sometimes these requirements are applied in a manner that appears to be a genuine attempt to determine how the particular kind of animal would have behaved in a specific type of scenario. This is perhaps most often true in situations where outside influences exclude the *sponte feritate* requirement. A good example is *Green v Naidoo*.⁶ A dog bit a child after the child had pulled a scab off its nose. Whereas biting would normally have been *contra naturam*, the child's unwitting pain-inflicting behaviour constituted "provocation" and "exonerated" the dog.⁷ However, in many other cases the *contra naturam* requirement is expressed in rigid, objective rules that, if anything, often appear to run contrary to the nature of the animals they pertain to, rather than to describe their behaviour realistically. Thus a dog that bites, a horse that kicks and an ox that butts, all behave *contra naturam*, unless one of the recognised defences are applicable,⁸ and such rules can be traced back to Roman law.⁹

It is significant that the requisite nature of the particular kind of animal is not determined by the evidence of people with expert knowledge of the behaviour and temperament of the relevant breed of animal. The court may listen to such evidence, but the court will itself determine whether the animal behaved *contra naturam*. Determining the nature of the animal's kind will in the final analysis be performed by people trained in the law, not by people trained in the natural sciences. This is very clearly illustrated in *Loriza Brahman v Dippenaar*.¹⁰ *In casu*, damage was caused by a Brahman calf. Expert evidence was led on the temperament and behaviour of the Brahman breed of cattle, to the effect that it is a nervous breed and that the kind of damage-inflicting behaviour that took place could be regarded as having been true to the nature of the Brahman breed, rather than contrary to its nature. The Supreme Court of Appeal did not agree, and found that the animal in question had acted *contra naturam*. In substantiating its decision to disregard the evidence concerning the nature of the Brahman breed, the court stated¹¹ that the nature of the *genus* had to be taken into account, not that of the *species*. Therefore, the question *in casu* was whether the Brahman calf had acted *contra* the nature of cattle, not the nature of the Brahman breed. In *Thysse v Bekker*¹² an expert opinion on the nature of a dog was also discounted in favour of an objective standard of behaviour imposed on the animal by the law, and given content by the judge.¹³ The court made it clear¹⁴ that consideration was not to be given to the nature of the specific animal, nor to the nature of its breed. The question was "whether the dog behaved in a manner which the law considers acceptable by animals which share the human environment with human beings because they have over the ages become domesticated".

Some judgments show an interesting trend to identify the *contra naturam* requirement with fault and/or wrongfulness as elements of delictual liability. In the

6 2007 6 SA 372 (W).

7 380–381. For other examples from case law cf Neethling and Potgieter 359; Van der Merwe 93.

8 Cf Neethling and Potgieter 358; Van der Merwe 92.

9 Van der Merwe *op cit*.

10 2002 2 SA 477 (SCA).

11 485H–J.

12 2007 3 SA 350 (SE) 354–355.

13 Cf Knobel "Contra naturam sui generis?" 2008 *THRHR* 496–498.

14 355A–C.

old judgment in *South African Railways v Edwards*,¹⁵ for instance, the then Appellate Division said:¹⁶ “[T]he idea underlying the *actio de pauperie*, an idea which is still at the root of the action, was to render the owner liable only in cases where so to speak the fault lay with the animal.” In a more recent example, *Green v Naidoo*¹⁷ made explicit mention of fault on the part of a dog. In no case is this kind of reasoning taken further than in *Da Silva v Otto*,¹⁸ where the court expected a dog to differentiate between a lawful and wrongful attack upon it. *In casu* the dog bit a person after the person had lawfully hit it, and therefore the dog behaved *contra naturam*. Explaining the applicable criterion, Nestadt J said¹⁹ that an objective test of the reasonable dog is used, and that a reasonable dog would submit to human authority and would not bite a person who launched a lawful attack on it. The similarity between this reasonable dog test and the reasonable person test of negligence speaks for itself.²⁰ Elsewhere in the same case the behaviour of a dog is furthermore said to be wrongful (“onregmatig”).²¹

In *Thysse v Bekker*,²² Jones J commented on the criteria formulated in *Da Silva*:

“This is by way of fiction, a convenient way of stating what the law expects of a domesticated animal. The courts do not literally attribute powers of reason to a dog by measuring its conduct against that of ‘the reasonable dog’, or literally ascribe to a dog the logical capacity to distinguish between an unjustified attack upon it which it may be ‘entitled’ to resist in a well-behaved manner and the lawful use of violence against it to restrain it, to which it should submit as a well-behaved domesticated animal.”

This *dictum* is in explicit agreement with Hunt:²³

“The *contra naturam* concept seems . . . to have come to connote ferocious conduct contrary to the gentle behaviour normally expected of domestic animals. This imports an objective standard suited to humans. It is far more refined than behaviour literally natural to that species of animal.”

In a somewhat less critical vein, Van der Walt and Midgley²⁴ state that the *contra naturam* requirement and the *sponte feritate* requirement are similar to the wrongfulness and fault enquiries respectively. They continue:

“*Prima facie* an attack by a domesticated animal is *contra naturam* and therefore unreasonable and wrongful. Yet this is not enough to found liability, for the animal must also have been at fault in behaving in such a way, in other words, motivated by its own instinct.”²⁵

On the strength of these *dicta* from case law and academic commentary, it should be clear that the *contra naturam* requirement makes blameworthiness of the animal pivotal to the liability of the owner of the animal. Irrespective of

15 1930 AD 3.

16 10.

17 2007 6 SA 372 (W) 378E–F.

18 1986 3 SA 538 (T).

19 542F–G.

20 Cf in general Neethling and Potgieter 131 *et seq*; Van der Walt and Midgley 166 *et seq* and case law cited.

21 541J.

22 2007 3 SA 350 (SE) 357A–B.

23 “Bad dogs” 1962 *SALJ* 328.

24 41.

25 *Ibid.*

whether the requirement is treated as an honest attempt to determine whether the behaviour of the animal was contrary to the nature of its kind, or whether the requirement is employed as a rigid set of behaviour rules to which domestic animals are subjected by the law, liability can only follow if the animal was in some way to blame for the harm. Of course, *dicta* from cases and commentaries that identify the *contra naturam* requirement with fault and/or wrongfulness on the part of the animal provide the clearest and most extreme evidence that a remnant of blameworthiness still lies at the root of the *actio de pauperie*.

3 REMNANTS OF BLAMEWORTHINESS IN DEFENCES AGAINST THE *ACTIO DE PAUPERIE*

Neethling and Potgieter point out that important defences against the *actio de pauperie* developed from the *sponte feritate* requirement: *vis maior*; culpable or provocative conduct on the part of the prejudiced person; culpable conduct on the part of a third person; and provocation by another animal. In addition, *volenti non fit iniuria* may also exclude liability.²⁶ A further defence is implied by the requirement that the prejudiced person or his or her property must be lawfully present at the place where and when the damage is inflicted by the animal. Thus, unlawful presence of the plaintiff or his or her property at the time constitutes a defence.²⁷ All these defences, with the possible exception of the *volenti* defence, strengthen the argument that remnants of blameworthiness are present in the *actio de pauperie*. *Vis maior* may be seen to remove blame from the animal, and therefore the owner is relieved from liability. The other defences may be seen to shift the blameworthiness away from the damage-causing animal to a person (either a third party or the prejudiced person himself) or to another animal. All these defences can thus be seen to strengthen a conclusion that some blameworthiness of the damage-causing animal is a prerequisite for holding the owner of the animal liable.

4 EVALUATION

If we accept that, at least at a subconscious level, some remnants of blameworthiness remain embedded in the substantive law relating to the *actio de pauperie*, the question that now presents itself is whether this should be so. Perhaps the most important thing to state first is that the outcome of South African cases dealing with the *actio* is, in general, satisfactory. Elsewhere, the present author has expressed an impression that reliance on the *actio de pauperie* may at times cause the potential applicability of other delictual actions to be overlooked, but in general, the outcome of case law does not offend one's sense of justice, and for this the courts must be applauded.²⁸ However, it is unacceptable to attach blameworthiness to an animal in our law, whether this happens explicitly or on a subconscious level. As pointed out above, an animal cannot understand the law, has no *locus standi*, and is not a legal subject in terms of our law. To attach blameworthiness to an animal is therefore illogical and unrealistic.

26 359; cf Van der Merwe 105 *et seq*; Van der Walt and Midgley 41 and case law cited.

27 Neethling and Potgieter 359 fn 47.

28 Knobel "Die *actio de pauperie* en aanspreeklikheid gebaseer op nalatigheid" 2010 *THRHR* 172.

The argument can be taken further. Blameworthiness has been used in a fairly loose sense in this contribution to signify a negative value judgment by the law. In a more precise sense, blameworthiness is, in the delictual sphere, usually associated with the element of fault.²⁹ It is trite in our law that to be at fault in the technical-legal sense, a person must be *culpa capax*, that is, he or she must be said to have capacity or be accountable. The test is that the person must be able to differentiate between right and wrong, and be able to act in accordance with that insight.³⁰ Can any animal be said to know the difference between right and wrong and to be able to act in accordance with this insight? An “obedient” dog may have learnt that “bad” behaviour on its part is visited with punishment while “good” behaviour may be rewarded with a pat, a friendly tone of voice or a snack. But does such an animal have the intellectual and moral faculties to know what is right and wrong? A human being under seven years of age is irrebuttably presumed to be *culpa incapax* and, when between seven and fourteen years, he or she is still presumed *culpa incapax*, albeit now rebuttably.³¹ How can an animal then be assumed to be accountable? Clearly, the answer must be that an animal should not be regarded as capable of being blameworthy in the eyes of the law.

As seen, an attempt was made in *Loriza* to refine the *contra naturam* requirement by stating that it referred to the *genus* and not the *species*. Again, this statement is problematic. The present author has pointed this out in a previous contribution.³² In zoology, a binomial system of naming animals is used. The first name refers to the genus and the second to the species. Thus, all modern breeds of cattle are thought to be derived from two species, namely *Bos taurus* and *Bos indicus*, or crosses between them. A reference to these names in *Loriza*³³ makes it tempting to conclude that the court had the zoological binomial names in mind when talking of genus and species. However, the genus *Bos* also includes wild cattle species such as the gaur, banteng and yak. The domestic dog shares its genus, *Canis*, with such species as wild wolves and jackals, and the domestic horse shares its genus, *Equus*, with such animals as wild zebras.³⁴ Clearly the zoological concept of genus cannot be used to determine the nature of domesticated animals.

If we accept that *genus* and *species* as used in *Loriza* must not be equated with the eponymous terms in zoology, counter-intuitive as this may be, a solution to the problem of determining the nature of the genus remains elusive. If genus must be taken to refer to “domestic cattle” or “domestic dog” or even “domestic animal” instead of different cattle or dog or other breeds, it remains impossible empirically to determine what the nature of such a genus is. How does one determine the nature of, say, a dog, if some dog breeds have been bred to be docile and phlegmatic and others to be fearless and ferocious attack dogs? The enormity of the problem is acknowledged by Satchwell J in *Green v Naidoo*,³⁵

29 Neethling and Potgieter 123; Van der Walt and Midgley 155.

30 Neethling and Potgieter 125–126; Van der Walt and Midgley 156–157.

31 Neethling and Potgieter 125; Van der Walt and Midgley 157.

32 Knobel 2010 *THRHR* 173–174.

33 2002 2 SA 477 (SCA) 485G–H.

34 Cf in general eg Grzimek (ed) *Grzimek's encyclopaedia of mammals* (1990) Vol 4 59 *et seq* 557 *et seq* Vol 5 386 *et seq*.

35 2007 6 SA 372 (W) 377F–H.

but is then side-stepped by asking what the nature of a “dog which is a household pet” would be. In respect of cattle, evidence was placed before the court in *Loriza*³⁶ that the two species of domesticated cattle differ appreciably in temperament. Clearly, if the nature of all domestic animals – expansive and inclusive as that concept is – must be considered, this cannot refer to some common denominator in the true, empirically-observed behaviour of these creatures in all their diversity, but must be an objective idea of how domesticated animals should, according to humans, behave.

The conclusion reached by the present writer on the *contra naturam* requirement is that, in the vast majority of applications, it can only function, as correctly pointed out by Jones J in *Thysse*, as a fiction or catch-phrase denoting a standard of behaviour imposed by the law on domestic animals, and one containing unacceptable remnants of blameworthiness at that.³⁷

5 AN ALTERNATIVE APPROACH

If having remnants of blameworthiness in the *actio de pauperie* is undesirable, the question arises whether a better alternative approach is available. The risk principle provides the best explanation why certain instances of delictual liability are strict rather than fault-based.³⁸ According to this principle, if a defendant creates an exceptional risk that others may be prejudiced, it may be desirable to visit that defendant with strict liability if such damage indeed materialises. This is certainly applicable to the keeping of domesticated animals. Experience has shown that keeping a domestic animal creates a higher than normal risk that others may be harmed. Therefore, even though the *actio de pauperie* may have arisen in a more primitive legal system to punish an animal for a perceived wrongdoing on the part of the animal, the existence of strict liability for damage caused by a domestic animal can in a modern legal system be justified with reference to the risk principle. One can conclude that a remnant of blameworthiness is not required to justify strict liability for the damage caused by a domestic animal. Strict liability for damage caused by animals is in step with international trends, and jurisdictions such as Belgium, England, France, Germany, the Netherlands, Spain and the United States of America all have at least some strict liability for damage caused by animals.³⁹

As we have seen, the remnants of blameworthiness play a pivotal part in the *actio de pauperie* in so far as they determine whether the owner of the animal will be liable or not. If having remnants of blameworthiness in the *actio* is undesirable, the next question that presents itself is what criteria should be used to determine whether the owner must be liable. As seen, Van der Walt and Midgley⁴⁰ compared the requirements of the action with the general elements of a delict. The requirements of the *actio de pauperie* are not usually expressed in terms of those elements. They are usually expressed, as was also done in this contribution, as a list of requirements that are unique to the action itself and that have become established since Roman times. Expressing the requirements in

36 2002 2 SA 477 (SCA) 485G–H.

37 Cf Knobel 2008 *THRHR* 498.

38 Cf Neethling and Potgieter 356; *Loriza* 485.

39 Koch and Koziol “Comparative conclusions” in Koch and Koziol (eds) *Unification of tort law: Strict liability* (2002) 396.

40 41.

terms of the standard delictual elements is problematic. First, and central to the theme of this contribution, it is inappropriate to attach legal value judgments like those involved in the elements of wrongfulness and fault to the behaviour of animals. Second, the present author has some doubts whether the element of 'conduct' or 'act' can be attributed to animals. An authoritative textbook on delict teaches that an "act" or "conduct" in the delictual sense is the act of a natural or juristic person, not an animal.⁴¹ Perhaps it would be prudent to talk of the behaviour of an animal rather than the "act" or "conduct" of the animal.

Be that as it may, and whether it is appropriate to refer to an "act" of an animal or not, there is no doubt that the behaviour of an animal can cause damage, and that the elements of damage and causation can therefore be attributed to the animal. In fact, this points to a fundamental difference between an animal and other dangerous things a person may own, such as, for instance, a motor car. A person creates a much higher risk that others will be prejudiced by owning a car than by owning an animal. However, a human act or omission is needed to set the car in motion, and if damage arises, that human conduct is the cause of the damage. An animal, on the other hand, can cause damage by itself, without human intervention. In the opinion of the present author, the element of causation may well hold the key to a rational and equitable application of the *actio de pauperie*.

In delict, factual and legal causation must both be present for liability to ensue, although, in general, legal causation is expressly considered only in difficult cases.⁴² Where a domestic animal is alleged to have caused damage, factual causation is, on the one hand, unlikely to cause particular difficulty and, on the other, is also unlikely to hold the key in differentiating between instances where the *actio de pauperie* must succeed and instances where it must fail. Using evidence and experience, and expressing the behaviour of the animal and the harm suffered by the plaintiff in the language of the *conditio sine qua non*, should quickly establish whether factual causation is present.⁴³

Legal causation is a different matter altogether. According to the celebrated judgment in *S v Mokgethi*,⁴⁴ the real enquiry is whether the link between the act and the damage is sufficiently close that the damage may be imputed to the wrongdoer taking account of policy considerations of reasonableness, fairness and justice.⁴⁵ Translated to the context of the *actio de pauperie*, legal causation is the enquiry whether the link between the behaviour of the animal and the plaintiff's damage is sufficiently close that the damage may be imputed to the animal, taking into account considerations of reasonableness, fairness and justice. An important aspect here is that 'imputed' does not imply blame, it refers to causation. If the damage is imputed to the animal, and this will happen if the damage was not too remote, the behaviour of the animal is legally deemed to be the cause of the damage. Policy considerations of reasonableness, fairness and justice play a part here, but not blameworthiness. Yardsticks and theories of legal causation,

41 Neethling and Potgieter 25.

42 Cf in general Neethling and Potgieter 175 *et seq*; Van der Walt and Midgley 196 *et seq*.

43 Cf in general Neethling and Potgieter 176–187; Van der Walt and Midgley 198–202 and case law cited.

44 1990 1 SA 32 (A) 40–41.

45 Cf also Neethling and Potgieter 191; *International Shipping Co (Pty) Ltd v Bentley* 1990 1 SA 680 (A) 700.

such as the presence of a *novus actus interveniens*, direct consequences and adequate causation can be used as subsidiary criteria to arrive at an answer.⁴⁶ It is noteworthy that, speaking of delict in general and not specifically the *actio de pauperie*, these criteria serve to endow conduct with a heightened legal relevance, but without necessarily implying that the conduct is legally reprehensible or that the defendant is legally blameworthy. Legal reprehensibility and blameworthiness are usually taken care of by the elements of wrongfulness and fault.⁴⁷ For this reason, criteria relating to wrongfulness and fault are not suited to be applied to the behaviour of animals, while the criteria relating to legal causation are most suitable for such an application.

At this stage, recognized defences against the action can play their part. Thus *vis maior*, or the conduct of a person, or the behaviour of another animal, may result in a finding that the behaviour of the animal was too remote and thus not the legal cause of the damage. Seen this way, and translated into the language of *Mokgethi*, one can say that it is reasonable, fair and just to hold the owner of an animal strictly liable if the animal caused the plaintiff's damage without being moved by a person or another animal to do so – not because the behaviour of the animal was blameworthy, but because its behaviour was the legal cause of the damage. Equally, it is not reasonable, fair and just to hold the owner liable if the animal was moved by a person or another animal to cause damage – not because the conduct of the animal has been cleared from blame, but because its conduct was too remote and cannot be regarded as the legal cause of the plaintiff's damage even though it may have been a factual cause thereof. This approach can steer clear from the unacceptable practice of expecting animals to behave with human qualities of morality and respect for the law.

In my opinion these insights enable us to place the *contra naturam* requirement into perspective. The *contra naturam* requirement, with its associated requirement that the animal must behave *sponte feritate commota*, should be seen as a formula to determine whether the behaviour of the animal was the legal cause of the plaintiff's damage. The best would actually be to drop the phrase *contra naturam sui generis* completely, and to reformulate the criterion to require that the domestic animal must have caused the damage of its own accord, and not because it has been moved by a person or another animal to behave in the manner it did.⁴⁸ The second best would be to acknowledge openly, as Jones J has already done in *Thysse v Bekker*, that the *contra naturam* requirement is a fiction in so far as it purports to endow animals with human qualities. All *dicta* susceptible to an interpretation that an animal is blameworthy must carefully be

46 *Mokgethi* 40–41; cf Neethling and Potgieter 191; Van der Walt and Midgley 203 and case law cited.

47 Cf in general, Neethling and Potgieter 33 *et seq* 123 *et seq*; Van der Walt and Midgley 67 *et seq* 155 *et seq* and case law cited. Recently, our courts stated that the element of wrongfulness is not concerned with the reasonableness of conduct, but rather with the reasonableness of holding someone delictually liable. The present author's respectful but firm view remains that whereas wrongfulness is undoubtedly concerned with the reasonableness of holding someone delictually liable, it is in essence also very intimately concerned with the question as to whether conduct is reasonable or unreasonable and hence legally reprehensible (Knobel "Thoughts on the functions and application of the elements of a delict" 2008 *THRHR* 650 *et seq*). This view is also held by eminent scholars in the field (see Neethling and Potgieter 78–82 for a detailed discussion).

48 Cf Knobel 2008 *THRHR* 499.

avoided. In doing this, the courts will continue to dispense justice to the prejudiced parties as they have been doing all along. In addition, they will also eradicate all signs that primitive remnants of blameworthiness have survived in the *actio de pauperie* of the present.

6 OTHER LEGAL SYSTEMS

In European legal codes with the most progressive approach to this matter, the owner and/or keeper, user or possessor of an animal is strictly liable for damage caused by the animal. This liability can be excluded by certain defences, but no enquiry into the nature of the animal is needed.⁴⁹ Examples are France,⁵⁰ the Netherlands,⁵¹ and Spain.⁵² Spanish commentators specifically state that the question whether the animal behaved *secundum naturam* or *contra naturam* makes no difference in Spanish law, and that this old distinction has been overcome.⁵³

Other codes may still wrestle with the dangerous “nature” or “characteristics” of animals. In England, the Animals Act of 1971 provides⁵⁴ that a keeper of an animal which does not belong to a dangerous species is liable for damage caused by that animal, if (a) the damage was of a kind the animal was likely to cause or if the damage was likely to be severe; (b) the likelihood of the occurrence or severity of the damage was due to characteristics of the animal that are not normally found in animals of that species; and (c) those characteristics were known to the keeper or certain other persons. Commentators point out that the rather complex provisions of the Act perpetuate the old *scienter* rule in terms of which strict liability for damage caused by a tame animal requires knowledge on the part of the defendant of an especially dangerous characteristic of that animal.⁵⁵ The *scienter* rule has on occasion been described as a primitive form of negligence on the part of the keeper of the animal.⁵⁶ In Germany, the form of liability depends on whether the animal is classified as a “useful” or “luxury” animal. Thus the keeper of an animal is strictly liable for damage caused by it, unless the animal is a “useful” or “working” animal used in the keeper’s profession or to provide for the keeper’s income, in which case liability is based on fault.⁵⁷ In a comparative analysis of strict liability for damage caused by animals under German and English law, Markesinis and Unberath⁵⁸ point out that in both jurisdictions strict liability is imposed only if the animals cause harm that can be described as a ‘typical’ consequence of the behaviour of that kind of animal.

49 Koch and Koziol “Comparative conclusions” in Koch and Koziol (eds) 397.

50 § 1385 CC; cf eg Galand-Carval “France” in Koch and Koziol (eds) 132–133.

51 § 6:179 NBW; cf eg Du Perron and Van Boom “Netherlands” in Koch and Koziol (eds) 238.

52 § 1905 CC; Martín-Casals, Ribot and Solé “Spain” in Koch and Koziol (eds) 298.

53 Martín-Casals, Ribot and Solé *op cit*, with reference to Domínguez *Responsabilidad civil extracontractual por daños causados por animales* (1997) 39.

54 S 2(2).

55 Rogers “England” in Koch and Koziol (eds) 111–112; Rogers *Winfield and Jolowicz on tort* (2010) 804–805.

56 Rogers in Koch and Koziol (eds) 112; Rogers 805 fn 34.

57 § 833 BGB; Fedtke and Magnus “Germany” in Koch and Koziol (eds) 154–155; Koch and Koziol in Koch and Koziol (eds) 397 fn 11.

58 *The German law of torts: A comparative treatise* (2002) 720–721.

This requirement is reminiscent of the notions underlying the “adequate cause” theory of legal causation.⁵⁹

It is submitted that the positions in these foreign jurisdictions vindicate the conclusions reached in this contribution. On the one hand, the most progressive legal provisions do not require any enquiry into the dangerous nature or characteristics of the animal that caused the damage. These trend-setting provisions probably point the way to the ideal position.⁶⁰ On the other hand, the relevant provisions in England and Germany attach importance to the dangerous nature or characteristics of animals, but do this to determine blameworthiness on the part of a person, not an animal. That such an enquiry is far more worthy of a modern legal system than the South African position should be self-evident.

7 CONCLUSION

The conclusions and recommendations reached in this contribution may be summarised as follows:

- (a) The *actio de pauperie* in South African law contains remnants of blameworthiness attached to animals. These remnants are unacceptable and should be removed from our law.
- (b) The best way of doing this would be to drop the *contra naturam*-requirement altogether. The second best way would be to acknowledge openly, as is done by Jones J in *Thysse v Bekker*, that the *contra naturam* requirement is a fiction in so far as it purports to endow animals with human qualities.
- (c) The requirement that the animal must behave spontaneously and not from influence by another animal or a person, should be interpreted to pertain to causation and not blameworthiness of the animal. Defences involving external influences by a person or another animal should also be interpreted to pertain to causation rather than blameworthiness.
- (d) The element of causation holds potential for a logical and equitable application of the *actio de pauperie*, especially criteria of legal causation that can endow behaviour with heightened legal relevance without branding it as blameworthy or legally reprehensible.

⁵⁹ Cf in general Neethling and Potgieter 193 *et seq.*

⁶⁰ Rogers 808.

The impact of the National Credit Act 34 of 2005 on standard acknowledgements of debt

Corlia van Heerden

BProc LLB LLMLLM LLD

Associate Professor of Mercantile Law, University of Pretoria

OPSOMMING

Die impak van die Nasionale Kredietwet op standaardkulderkennings

Die Nasionale Kredietwet 34 van 2005 (die NKW of Wet) is op 'n veel wyer reeks kredietooreenkomste van toepassing as sy voorganger, die Wet op Kredietooreenkomste 75 van 1980, wat net op afbetalingsverkoopsooreenkomste en huurooreenkomste van toepassing was, en is ook nie soos voormelde wet aan 'n monetêre plafon onderhewig nie. Die omvattende beskerming wat die NKW bied word verder uitgebrei deurdat artikel 8(4)(f) van die Wet voorsiening maak dat benewens die benoemde krediettransaksies gelys in artikel 8(4)(a) tot (e) van die Wet, enige ander ooreenkoms, anders as 'n kredietfasiliteit of kredietwaarborg, waarkragtens betaling van 'n bedrag verskuldig deur een persoon aan 'n ander uitgestel word en enige heffing, fooi of rente betaalbaar is aan die kredietverskaffer ten aansien van die ooreenkoms of die bedrag wat uitgestel is, as 'n krediettransaksie vir doeleindes van die NKW kwalifiseer. In die praktyk gebeur dit dikwels dat 'n skuldeiser aan wie 'n skuldenaar 'n bedrag geld verskuldig is op grond van 'n bestaande skuld, vanweë die skuldenaar se onvermoë om die skuld te betaal op die datum soos aanvanklik ooreengekom, die skuldenaar akkommodeer deur 'n standaardkulderkenning met hom aan te gaan ten einde afbetaling van die skuld in paaielemente en teen heffing van rente en ander koste te reël. Die doel van hierdie bydrae is om vas te stel of 'n standaardkulderkenning aangegaan op of na 1 Junie 2007 wel as 'n artikel 8(4)(f) krediettransaksie kwalifiseer, die gevolge van toepassing van die NKW op sodanige kulderkenning uit te wys en voorstelle te maak om sodanige kulderkenning buite die toepassingsgebied van die Wet te plaas.

1 INTRODUCTION

The National Credit Act¹ (“the NCA” or “Act”) has provided South Africans with a comprehensive credit act which governs an extended range of credit agreements. This is due to the fact that the Act not only regulates instalment sale agreements and lease agreements in respect of movables as was done by its predecessor – the repealed Credit Agreements Act² – but applies to a much wider variety of credit agreements and also does not cap credit agreements at a specific monetary ceiling as was done by the Credit Agreements Act which donned a monetary regulatory cap of R500 000. As indicated below, credit agreements as

1 34 of 2005, which came into operation on 1 June 2007.

2 75 of 1980.

provided for in the NCA usually entail some deferral of payment accompanied by interest and/or a fee or charge.

In practice it often happens that a creditor accommodates a debtor by entering into an acknowledgement of debt with the debtor to facilitate repayment of an existing debt instead of instituting legal action for recovery of the debt. Thus, for instance, where a creditor lends an amount of money to a debtor which has to be repaid by a certain date, it will often happen that the debtor is unable to repay the money on the due date and the creditor would then accommodate the debtor by entering into an acknowledgement of debt with the debtor as a result of which the debtor will be afforded a longer period to repay the debt. The standard acknowledgement of debt usually provides for a deferral of payment in the form of instalments and further provides for interest to be charged as well as fees such as legal fees and collection commission. The question therefore arises whether a standard acknowledgment of debt constitutes a credit agreement for purposes of the NCA.

This discussion focuses on the application of the NCA to various types of credit agreements and whether it can be said that a standard acknowledgement of debt constitutes a credit agreement that falls within the ambit of the Act. The consequences, should the Act find application to standard acknowledgements of debt, are also discussed and suggestions are made regarding possible ways in which the issue may be addressed. It should be noted that this discussion is limited to standard acknowledgements of debt entered into on or after 1 June 2007.

2 SCOPE AND APPLICATION OF NCA

2.1 General

The NCA applies to every credit agreement between parties dealing at arm's length and made within or having an effect within, the Republic of South Africa.³ "Credit", when used as a noun, is defined in the Act as a deferral of payment of money owed to a person or a promise to defer such payment; *or* a promise to advance or pay money to or at the direction of another person.⁴ "Agreement" includes an arrangement or understanding between or among two or more parties which purports to establish a relationship in law between those parties.⁵ Thus it is essential to establish whether a specific agreement entered into in South Africa or having an effect in South Africa constitutes a credit agreement as provided for in the NCA. It should also be borne in mind that even if an agreement constitutes a credit agreement as envisaged by the NCA, the Act will not apply if the agreement was not concluded at arm's length.

3 S 4(1). This application is subject to s 5 (which provides for the application of the NCA to incidental credit agreements) and s 6 (which provides for the limited application of the NCA to certain juristic persons). See further Otto and Otto *The National Credit explained* (2010) ch 3 and Stoop "Kritiese evaluasie van die toepassingsveld van die 'National Credit Act'" 2008 *De Jure* 352. (All references to sections hereinafter are to sections of the NCA, unless otherwise indicated.)

4 S 1.

5 S 1.

On the topic of arm's length, the NCA specifically provides that in any of the following arrangements the parties are not dealing at arm's length:⁶

- (a) A shareholder loan or other credit agreement between a juristic person, as consumer, and a person who has a controlling interest in that juristic person, as credit provider;
- (b) a loan to a shareholder or other credit agreement between a juristic person, as credit provider, and a person who has a controlling interest in that juristic person, as consumer;
- (c) a credit agreement between natural persons who are in a familial relationship and are co-dependent on each other or one is dependent upon the other; and
- (d) any other arrangement in which a party is not independent of the other and consequently does not necessarily strive to obtain the utmost possible advantage out of the transaction; or that is of a type that has been held in law to be between parties who are not dealing at arm's length.

2.2 Parties to a credit agreement

The parties to a credit agreement governed by the NCA are referred to as the "consumer" and the "credit provider". A "consumer" in respect of a credit agreement to which the NCA applies, means⁷

- (a) the party to whom goods or services are sold under a discount transaction, incidental credit agreement or instalment agreement;
- (b) the party to whom money is paid, or credit granted, under a pawn transaction;
- (c) the party to whom credit is granted under a credit facility;
- (d) the mortgagor under a mortgage agreement;
- (e) the borrower under a secured loan;
- (f) the lessee under a lease;
- (g) the guarantor under a credit guarantee; or
- (h) the party to whom or at whose direction money is advanced or credit granted under any other credit agreement.

A "credit provider" in respect of any credit agreement to which the NCA applies, means⁸

- (a) the party who supplies goods or services under a discount transaction, incidental credit agreement or instalment agreement;
- (b) the party who advances money or credit under a pawn transaction;
- (c) the party who extends credit under a credit facility;
- (d) the mortgagee under a mortgage agreement;
- (e) the lender under a secured loan;
- (f) the lessor under a lease;

⁶ S 4(2)(b). See also *Beets v Swanepoel* [2010] JOL 26422 (NC).

⁷ S 1.

⁸ S 1.

- (g) the party to whom an assurance or promise is made under a credit guarantee;
- (h) the party who advances money or credit to another under any other credit agreement; or
- (i) any other person who acquires the rights of a credit provider under a credit agreement after it has been entered into.

2 3 Types of credit agreements

Three main types of credit agreements are regulated by the NCA, namely, credit facilities, credit transactions and credit guarantees or a combination of any of the aforesaid credit agreements.⁹ It is expressly provided that a policy of insurance or credit extended by an insurer solely to maintain the payment of premiums on a policy of insurance, and a lease of immovable property or a transaction between a *stokvel* and a member of that *stokvel* in accordance with the rules of that *stokvel*, irrespective of their form, are not credit agreements for purposes of the NCA.¹⁰

2 3 1 Credit facility

An agreement, irrespective of its form, but not including an agreement contemplated in section 8(2) or section 8(4)(6)(b), constitutes a credit facility if in terms of that agreement¹¹

- (a) a credit provider undertakes to supply goods or services or to pay an amount or amounts, as determined by the consumer from time to time, to the consumer or on behalf of, or at the direction of, the consumer, and either to defer the consumer's obligation to pay any part of the cost of goods or services, or to repay to the credit provider any part of an amount, or bill the consumer periodically for any part of the cost of goods or services, or any part of an amount; and
- (b) any charge, fee or interest is payable to the credit provider in respect of any amount deferred or any amount billed and not paid within the time provided for in the agreement.

2 3 2 Credit transaction

An agreement, irrespective of its form, but not including an agreement contemplated in section 8(2) as indicated above, constitutes a credit transaction if it is¹²

- (a) a pawn transaction¹³ or discount transaction;¹⁴

⁹ S 8(1)(a).

¹⁰ S 8(2).

¹¹ S 8(3).

¹² S 8(4).

¹³ A pawn transaction means an agreement, irrespective of its form, in terms of which

(a) one party advances money or grants credit to another, and at the time of doing so, takes possession of goods as security for the money advanced or credit granted; and

(b) either

(i) the estimated resale value of the goods exceeds the value of the money provided or the credit granted; or

(ii) a charge, fee or interest is imposed in respect of the agreement, or in respect of the amount loaned or credit granted; and

- (b) an incidental credit agreement¹⁵ subject to section 5(2);¹⁶
- (c) an instalment agreement;¹⁷
- (d) a mortgage agreement¹⁸ or secured loan;¹⁹
- (e) a lease;²⁰ or

(c) the party that advanced the money or granted the credit is entitled on expiry of a defined period to sell the goods and retain all the proceeds of the sale in settlement of the consumer's obligations under the agreement.

- 14 A discount transaction means a transaction, irrespective of its form, in terms of which (a) goods or services are to be provided to a consumer over a period of time; and (b) more than one price is quoted for the goods or service, the lower price being applicable if the account is paid on or before a determined date, and a higher price or prices being applicable if the price is paid after that date, or is paid periodically during the period.
- 15 It is to be noted that the NCA has limited application to an incidental credit agreement as set out in s 5. An incidental credit agreement means an agreement, irrespective of its form, in terms of which an account was tendered for goods or services that have been provided to the consumer, or goods or services that are to be provided to a consumer over a period of time and either or both of the following conditions apply: (a) a fee, charge or interest became payable when payment of an amount charged in terms of that account was not made on or before a determined period or date; or (b) two prices were quoted for the settlement of the account, the lower price being applicable if the account is paid on or before a determined date, and the higher price being applicable due to the account not having been paid by that date.
- 16 In terms of s 5(2) the parties to an incidental credit agreement are deemed to have made that agreement on the date that is 20 business days after the supplier of the goods or services that are the subject of that account, first charges a late payment fee or interest in respect of that account; or a pre-determined higher price for full settlement of the account first becomes applicable.
- 17 An instalment agreement means a sale of movable property in terms of which
- (a) all or part of the price is deferred and is to be paid by periodic payments;
 - (b) possession and use of the property is transferred to the consumer;
 - (c) ownership of the property either
 - (i) passes to the consumer only when the agreement is fully complied with; or
 - (ii) passes to the consumer immediately subject to a right of the credit provider to repossess the property if the consumer fails to satisfy all of the consumer's financial obligations under the agreement; and
 - (d) interest, fees or other charges are payable to the credit provider in respect of the agreement, or the amount that has been deferred.
- 18 A mortgage means a pledge of immovable property that serves as security for a mortgage agreement.
- 19 A secured loan means an agreement, irrespective of its form but not including an instalment agreement, in terms of which a person (a) advances money or grants credit to another; and (b) retains, or receives a pledge or cession of the title to any movable property or other thing of value as security for all amounts due under that agreement.
- 20 A lease means an agreement in terms of which
- (a) temporary possession of any movable property is delivered to or at the direction of the consumer, or the right to use any such property is granted to or at the direction of the consumer;
 - (b) payment for the possession or use of that property is
 - (i) made on an agreed or determined periodic basis during the life of the agreement; or
 - (ii) deferred in whole or in part for any period during the life of the agreement;
 - (c) interest, fees or other charges are payable to the credit provider in respect of the agreement, or the amount that has been deferred; and

continued on next page

- (f) any other agreement, other than a credit facility or credit guarantee, in terms of which payment of an amount owed by one person to another is deferred, and any charge, fee or interest is payable to the credit provider in respect of the agreement or the amount that has been deferred.

2 3 3 *Credit guarantee*

An agreement, irrespective of its form, but not including an agreement contemplated in section 8(2) as indicated above, constitutes a credit guarantee if, in terms of that agreement, a person undertakes or promises to satisfy upon demand any obligation of another consumer in terms of a credit facility or a credit transaction to which the NCA applies.²¹

2 3 4 *Combined credit agreement*

- (a) If a particular credit agreement constitutes both a credit facility and a credit transaction,
- (b) subject to (c) below, that agreement is equally subject to any provision of the NCA that applies specifically or exclusively to either credit facilities or mortgage agreements or secured loans, as the case may be; and
- (c) for the purpose of applying section 108,²² that agreement must be regarded as a credit facility; or in terms of section 4(1)(b)²³ read with section 9(4),²⁴ that agreement must be regarded as a large agreement if it is a mortgage agreement.

2 4 Further categories of credit agreements based on size of agreement

Credit agreements are further divided into small, large and intermediate credit agreements.²⁵ This distinction influences a number of aspects such as disclosure requirements and exemption from the application of the Act when entered into by certain juristic persons.

A credit agreement is a small agreement if it is a pawn transaction; a credit facility with a limit falling at or below R15 000 and any other credit transaction, except a mortgage agreement or credit guarantee, and the principal debt under that transaction or guarantee falls at or below R15 000.²⁶ It qualifies as an intermediate agreement if it is a credit facility with a credit limit that falls above R15 000 or any credit transaction except a pawn transaction, a mortgage agreement or a credit guarantee, and the principal debt under that transaction or guarantee falls between R15 000 and R250 000.²⁷ Large credit agreements are

- (d) at the end of the term of the agreement, ownership of that property either
- (i) passes to the consumer absolutely; or
 - (ii) passes to the consumer upon satisfaction of specific conditions set out in the agreement.

21 S 8(5).

22 S 108 deals with statements of account.

23 S 4(1)(b) provides for exemption from the application of the NCA where a small juristic person with an asset value or annual turnover of less than R1 million enters into a small or intermediate credit agreement.

24 S 9(4) indicates when a credit agreement will constitute a large agreement for purposes of the NCA.

25 S 9(1).

26 S 9(2).

27 S 9(3).

mortgage agreements or any other credit transaction except a pawn transaction or a credit guarantee, if the principal debt under that transaction or guarantee falls at or above R250 000²⁸.

Pawn agreements will thus always be small agreements and mortgage agreements will always be large agreements whilst it should be noted that credit facilities can be small or intermediate agreements but are not treated as large agreements even though they may fall within the monetary threshold for large agreements.

3 EXEMPT AGREEMENTS AND LIMITED APPLICATION OF NCA TO JURISTIC PERSONS

An investigation into the scope and application of the NCA and the question as to whether a standard acknowledgement of debt constitutes a credit agreement will be incomplete unless the exemptions to the application of the Act and the instances of limited application of the Act to juristic persons are considered.

For purposes of comprehending the extent of the exemptions relating to juristic persons it is appropriate to point out that the NCA contains an extended definition of a juristic person which includes a partnership, association or other body of persons, corporate or unincorporated, or a trust if there are three or more individual trustees or the trustee itself is a juristic person, but does not include a *stokvel*.²⁹

3.1 Exempt agreements

The following agreements are expressly provided to be exempt from the application of the Act, namely:

- (a) a credit agreement in terms of which the consumer is
 - (i) a juristic person whose asset value or annual turnover, *together with the combined asset value or annual turnover of all related*³⁰ *juristic persons* (my emphasis) at the time the agreement is made, equals or exceeds R1 million³¹ (hereinafter a “large” juristic person);
 - (ii) the State;³² or
 - (iii) an organ of State;³³
- (b) a large agreement in terms of which the consumer is a juristic person whose asset value or annual turnover is, at the time the agreement is made, below R1 million³⁴ (hereinafter “a small juristic person”);
- (c) a credit agreement in terms of which the credit provider is the Reserve Bank of South Africa;³⁵ or

28 S 9(4).

29 S 1. It should thus be noted that a trust will consequently qualify as a natural person if it has less than three trustees.

30 According to s 4(2)(d), a juristic person is related to another juristic person if one of them has direct or indirect control over the whole or part of the business of the other; or a person has direct or indirect control over both of them.

31 S 4(1)(a)(i).

32 S 4(1)(a)(ii). The “state” is not defined in the NCA.

33 S 4(1)(a)(iii).

34 S 4(1)(b). It should be noted that the asset value or annual turnover of related juristic persons are not taken into account for purposes of this specific exemption.

- (d) a credit agreement in respect of which the credit provider is located outside the Republic, approved by the Minister on application by the consumer in the prescribed manner and form.³⁶

Note should also be taken of section 4(2)(c) which provides that the NCA applies to a credit guarantee only to the extent that the Act applies to a credit facility or credit transaction in respect of which the credit guarantee is granted. This effectively means that if the credit transaction or credit facility in respect of which the credit guarantee is granted falls outside the application of the NCA, the Act will also not apply to the credit guarantee. Thus the surety will not be able to rely on the provisions of the NCA for protection. As the subsection provides that the Act applies only to a credit guarantee *to the extent* (my emphasis) that the Act applies to the credit facility or credit transaction in respect of which the credit guarantee is granted, it is further submitted that a natural person consumer who stood surety for a juristic person to whom the Act applies, for example in respect of an intermediate credit transaction entered into by a small juristic person, will not be able to rely on the provisions of the Act relating to reckless credit and over-indebtedness as those provisions do not apply to juristic persons and the surety will only be afforded the protection of the Act *to the extent* that the Act applies to the underlying agreement.

It is further provided that the application of the NCA extends to a credit agreement or proposed credit agreement irrespective of whether the credit provider resides or has its principal office within or outside the Republic; or (subject to section 4(1)(c)) is an organ of state; an entity controlled by an organ of state; or an entity created by any public regulation or the Land and Agricultural Development Bank.³⁷ In terms of section 4(4) of the Act, if the NCA applies to a credit agreement it continues to apply to that agreement even if a party to that agreement ceases to reside or have its principal office within the Republic and it applies in relation to every transaction, act or omission under that agreement, whether that transaction, act or omission occurs within or outside the Republic.³⁸

3 2 Limited application of NCA to juristic persons

From the aforementioned it is thus clear that, whereas the Act will not apply to credit agreements entered into with “large” juristic persons as indicated in section 4(1)(a)(i) or to large credit agreements, such as mortgage bonds, entered into by “small” juristic person consumers, the Act will in fact apply to small and intermediate credit agreements entered into by small juristic person consumers. However, such application is limited by section 6 of the NCA which indicates that the following provisions of the Act do not apply to a credit agreement or proposed credit agreement in terms of which the consumer is a juristic person:

- (a) Chapter 4, Parts C and D, which deal with credit marketing practices and over-indebtedness³⁹ and reckless credit⁴⁰ respectively;

35 S 4(1)(c).

36 S 4(1)(d). See reg 2 for the prescribed manner and form 1 for the prescribed form.

37 S 4(3)(a) and (b).

38 S 4(4)(a) and (b). See further s 4(5) for “exemptions” relating to cheques and charges against credit facilities.

39 See s 79.

40 See s 80.

- (b) Chapter 5, Part A, section 89(2)(b) which deals with an agreement resulting from negative option marketing;
- (c) Chapter 5, Part A, section 90(2)(o) which deals with agreements at a variable interest rate; and
- (d) Chapter 5, Part C, which deals with the consumer's liability, interest, charges and fees.

Thus it may be concluded that with the exception of the parts and provisions of the NCA stated above, the remainder of the provisions of the NCA apply to those small juristic persons who enter into small and intermediate credit agreements.

As a result of the aforesaid limited application of the NCA to small juristic persons who enter into small and intermediate credit agreements, such juristic persons enjoy considerably less benefits under the NCA and will for instance not be able to access the debt relief provisions of the Act in respect of reckless credit and over-indebtedness and will also not be able to rely on the protection the Act affords consumers by prescribing certain rates of interest⁴¹ or by imposing the statutory *in duplum* rule contained in section 103(5).

4 DOES A STANDARD ACKNOWLEDGEMENT OF DEBT CONSTITUTE A CREDIT AGREEMENT FOR PURPOSES OF THE NCA?

From the aforementioned explanation regarding the scope and application of the NCA it may be observed that a standard acknowledgement of debt entered into to arrange the repayment terms of an existing debt does not constitute a credit facility as it does not entail an undertaking by a credit provider to provide goods or services to or at the direction of the consumer – the goods or services have already been rendered sometime in the past as a result of another agreement or cause of action (such as a delict) between the parties. The purpose of the acknowledgement of debt is merely to negotiate how the amount owing to the credit provider is going to be repaid. These terms of repayment usually differ from the terms of repayment applicable to the original agreement between the parties. It is also clear that a standard acknowledgement of debt does not constitute a credit guarantee or any of the named credit transactions such as a pawn agreement, discount agreement, incidental credit agreement, instalment agreement, lease, secured loan or mortgage agreement.

However, due to the fact that it contains a deferral of payment and requires the payment of interest, fees and other charges as a result of such deferral, a standard acknowledgement of debt as described herein appears to fall within the ambit of the “catch all” credit transaction provided for in section 8(4)(f) as indicated above. This was indeed what was *inter alia* decided in *Carter Trading v Blignaut*,⁴² where the court unfortunately did not elaborate on the implications of its

⁴¹ See reg 42.

⁴² 2010 2 SA 46 (ECP). At para 14 it is indicated that the acknowledgement *inter alia* contained the following clause: “1. I undertake to pay the undermentioned amount, interest calculated monthly in advance from . . . on the balance of the capital owing from time to time at the rate of 15.5% [fifteen and a half percent] per annum, the cost of preparing and negotiating and preparing this acknowledgement of debt and collection commission calculated with the Rules of the Law Society of the Cape of Good Hope.” The court held (paras 16 and 17) that in its view the payment of the amount owing was deferred to 24 December

findings other than to indicate that it requires compliance with the pre-enforcement provisions contained in sections 129 and 130 of the NCA. In *Grainco (Ltd) v Broodryk*,⁴³ however, the court after mentioning that the defendants alleged that the acknowledgement of debt constituted a section 8(4)(f) credit agreement and thus necessitated compliance with sections 129 and 130 of the Act, indicated that the underlying cause of action in respect of which the acknowledgement of debt *in casu* was entered into was a damages claim and not a money lending transaction.⁴⁴ Cillie J remarked that he agreed with counsel for the plaintiff that it could never have been the intention of the legislature that such a transaction should fall within the field of application of the NCA as it would constitute “an absurdity so glaring it could never have been contemplated by the legislature”.⁴⁵ The judge did not deal with the issue any further, save to indicate that the preamble to the NCA in which the objectives of the Act are described, confirms that it was not the objective that the Act should apply to such a mutual postponement of payment of damages.⁴⁶ However, it is submitted that it is unclear exactly which wording in the preamble to the Act prompted the court to come to the aforesaid conclusion.⁴⁷ It is further to be noted that in neither of the aforementioned cases the consequences of regarding an acknowledgment of debt as a section 8(4)(f) credit agreement were considered, save the indication in

2008 and that the defendant undertook to pay, in addition to the amount owing, at least the cost of preparing the acknowledgement of debt and in the event of a failure to pay the sum owing, also collection commission and legal fees and that it would appear that those terms are exactly what is envisaged in the Act to be a s 8(4)(f) credit agreement. Van der Byl AJ remarked that there were other and perhaps even more persuasive considerations on which the acknowledgement in question must be adjudged as being a credit agreement envisaged in the Act (para 19). He then quoted the definition of a credit facility as per the NCA and indicated that insofar as the plaintiff provided goods to the defendant on credit on the basis of the acknowledgement of debt which was eventually concluded, it would appear that such an agreement would in any event have been a credit agreement, ie a credit facility (paras 21–23). In response to the defendant’s submission that an acknowledgement of debt constitutes a settlement between the parties and therefore a novation of the defendant’s obligation to pay for the goods sold and delivered, the court indicated that the acknowledgement of debt in the present matter was not a novation of the obligations of the defendant under the agreement in respect of the goods sold and delivered but rather appeared to be a confirmation that created a further obligation relating to the same performance and not as a replacement of the existing obligation (para 25).

43 2011 JDR 0172 (FB).

44 Para 7.3.

45 Para 7.4.

46 Para 7.5. It should be noted that the court did not indicate on which word or sentence in the preamble it based this conclusion.

47 The preamble reads as follows: “To promote a fair and non-discriminatory marketplace for access to consumer credit and for that purpose to provide for the general regulation of consumer credit and improved standards of consumer information; to promote black economic empowerment and ownership within the consumer credit industry; to prohibit certain unfair credit and credit marketing practices; to promote responsible credit granting and use and for that purpose to prohibit reckless credit granting; to provide for debt reorganization in cases of over-indebtedness; to regulate credit information; to provide for registration of credit bureaux, credit providers and debt counselling services; to establish national norms and standards relating to consumer credit; to provide a consistent enforcement framework relating to consumer credit; to establish the National Credit Regulator and the National Consumer Tribunal; to repeal the Usury Act, 1968 and the Credit Agreements Act, 1980; and to provide for related incidental matters.”

Carter v Blignaut that it would require compliance with sections 129 and 130 of the Act.

The question therefore arises whether the legislature could ever have intended that a standard acknowledgement of debt, entered into on or after 1 June 2007, for purposes of rearranging the repayment terms of an existing debt in respect of which for instance money has already been advanced to a consumer a considerable period of time ago or where damages were suffered as a result of a delict such as a motor vehicle accident, should constitute a credit agreement for purposes of the NCA. Unfortunately the Act contains no clear indication regarding the intention of the legislature. Section 2(1) provides that the Act must be interpreted in a manner that gives effect to the purposes set out in section 3. The latter section contains no indication that the legislature intended standard acknowledgements of debt relating to the repayment of an existing debt to be excluded from the application of the Act and, it is submitted, nor does the section contain any provision from which such a specific intention may be inferred. If the legislature had intended to exclude acknowledgements of debt from the application of the Act, one would have expected it to be mentioned together with the exclusions in respect of insurance policies, leases of immovable property and *stokvels* provided for in section 8(2) or at least for it to be mentioned as an exemption under section 4. Standard acknowledgements of debt of the nature discussed herein are further usually entered into on terms that are in the best interests of the credit provider who is already displaying some leniency towards the consumer and would thus also not be able to escape the application of the Act on the basis that they are not entered into at arm's length. Due to the elements of deferral and charging of interest, fees and other charges in a standard acknowledgement of debt and in the absence of any express or implicit indication to the contrary, it would thus seem an inescapable conclusion that standard acknowledgements of debt entered into with natural person consumers and small juristic persons who conclude small or intermediate credit agreements fall within the category of credit transactions envisaged by section 8(4)(f) of the NCA.

5 CONSEQUENCES OF TREATING AN ACKNOWLEDGEMENT OF DEBT AS A CREDIT AGREEMENT

5.1 Where the acknowledgement of debt is entered into with a natural person

Acknowledgements of debt are usually entered into with natural persons even though the initial debtor in respect of the original debt or agreement in respect of which the acknowledgement of debt is signed, may have been a juristic person.

Regarding an acknowledgement of debt entered into with a natural person consumer on or after 1 June 2007 as a section 8(4)(f) credit transaction, it means that the creditor will have to comply with the NCA in all respects. It will *inter alia* have the following serious implications:

- (a) Where the creditor, alone or in conjunction with any associated person,⁴⁸ is the credit provider under at least 100 credit agreements (other than incidental

⁴⁸ According to s 40(2)(d), an "associated person" with respect to a creditor who is a natural person includes the credit provider's spouse or business partners. With respect to a credit provider who is a juristic person, it includes any person that directly or indirectly has a

credit agreements) or where the total principal debt owing to that creditor under all outstanding credit agreements (other than incidental credit agreements) exceeds R500 000, the creditor will have to apply to be registered as a credit provider.⁴⁹ Failure to register as a credit provider whilst under the obligation to register as such has the serious consequence that a credit agreement entered into by a creditor who is required to be registered as a credit provider under the NCA, but who is not so registered, is unlawful.⁵⁰ If a credit agreement is unlawful as envisaged in section 89, a court must order that the agreement is void as from the date it was entered into. It must then further order that the credit provider must refund to the consumer any money paid by the consumer under that agreement to the credit provider together with interest.⁵¹ In addition, all the purported rights of the credit provider under that agreement to recover any money paid or goods delivered to, or on behalf of, the consumer, in terms of that agreement, are either cancelled, unless the court concludes that doing so in the circumstances would unjustly enrich the consumer; or forfeited to the state, if the court concludes that cancelling those rights in the circumstances would unjustly enrich the consumer.⁵² Thus, clearly a no-win situation for the creditor.

- (b) Prior to entering into the acknowledgement of debt, the creditor will be obliged to conduct an extensive assessment as provided for in section 81 of the Act. Where he fails to do so the agreement will *per se* constitute reckless credit. Even where an assessment is duly conducted prior to entering into the acknowledgement of debt, it would still be possible for the consumer to raise reckless credit on the basis that, despite the assessment, he did not understand his risks, costs and obligations under the acknowledgement of debt or that entering into the acknowledgement of debt made him over-indebted as envisaged by the NCA.⁵³
- (c) A debtor would be able to aver that he has become over-indebted since entering into the acknowledgement of debt, for example because he was retrenched. It will then be possible for him to voluntarily apply for debt review in terms of section 86 of the NCA in order to obtain a moratorium on enforcement⁵⁴ and achieve a restructuring of his credit agreement debt,⁵⁵ which will then include the debt as provided for in the acknowledgement of debt. If unable to apply for voluntary debt review prior to institution of action, the debtor would be able once court proceedings have commenced, to

controlling interest in the credit provider, or is directly or indirectly controlled by the credit provider; any person that is directly or indirectly controlled by a person contemplated as aforesaid or any credit provider that is a joint venture partner of a person contemplated in s 47(2)(d).

49 S 40(1)(a) and (b).

50 S 40(4) read with s 89(2)(d).

51 S 89(5)(b). The interest is calculated at the rate set out in the agreement and for the period from the date on which the consumer paid the money to the credit provider, until the date the money is refunded to the consumer.

52 S 89(5)(c)(i) and (ii). See *Cherangani Trade and Investment 107 (Edms) Bpk v Mason* unreported case no 6712/2008 (O).

53 S 80 read with s 79.

54 S 88(3). See further Scholtz *et al Guide to the National Credit Act* (2008) para 11.3.3.6.

55 S 86(7)(c) read with s 87.

apply to the court in its discretion to refer him to a debt counsellor in accordance with section 85 of the Act.⁵⁶

- (d) The provisions of the NCA regarding interest and costs as set out in sections 100 to 107, read with regulations 39 to 48, apply to the agreement. This *inter alia* means that the creditor will not be able to recover more interest than the prescribed rate as set out in regulation 42 and also that the statutory *in duplum* rule as set out in section 103(5)⁵⁷ operates in favour of the debtor.
- (e) The consumer rights relating *inter alia* to the right to receive documents in an official language, the use of plain and understandable language and the right to receive documents as set out in Part A⁵⁸ of Chapter 4 of the Act as well as the provisions of the Act relating to confidentiality, personal information and consumer credit records will apply to the agreement.
- (f) The provisions of the Act regarding unlawful agreements, provisions and supplementary agreements as set out in sections 89, 90 and 91 respectively, will apply to the agreement and the debtor will most probably also be entitled to statements of account as provided for in the NCA.⁵⁹
- (g) Before any legal action can be taken against the consumer, the creditor will have to comply with the provisions of Part C of Chapter 6 of the Act relating to pre-enforcement procedures. This will entail affording the consumer a notice in terms of either section 129(1)(a)⁶⁰ (or 86(10)⁶¹ should the debtor be under debt review).
- (h) Once legal action is instituted the creditor will have to allege and observe extensive procedural compliance as set out in section 130(1) and (3) of the Act, which if not complied with will provide the debtor with a variety of technical defences and objections that he would not otherwise have had.
- (i) The court will further be provided with an extended range of powers as catered for in section 130(4).⁶²
- (j) It is debatable whether the agreement will also have to comply with the disclosure and content requirements of the Act as it seems illogical and absurd to require such compliance, including providing a quotation, given the nature and circumstances of the acknowledgement of debt.

⁵⁶ See Scholtz *et al* para 11.3.3.5 for a detailed discussion of s 85 of the Act.

⁵⁷ S 103(5) provides that despite any provision of the common law or a credit agreement to the contrary, the amounts contemplated in s 101(1)(b)–(g) that accrue during the time that a consumer is in default under the credit agreement may not, in aggregate, exceed the unpaid balance of the principal debt under that credit agreement as at the time that the default occurs. As regards the operation of the statutory *in duplum* rule, see *Nedbank Ltd v National Credit Regulator* 2011 3 SA 581 (SCA).

⁵⁸ See ss 60–66.

⁵⁹ See ss 107–115.

⁶⁰ S 129(1)(a) obliges the credit provider to deliver a written notice to the consumer-debtor in which the consumer's attention is drawn to his default under the credit agreement and it is proposed that the consumer refer the credit agreement to a debt counsellor, alternative dispute resolution agent, consumer court or ombud with jurisdiction, with the intention that the parties resolve any dispute under the agreement or develop and agree on a plan to bring the payments under the agreement up to date.

⁶¹ S 86(10) provides for termination of debt review where the consumer is in default and at least 60 business days have lapsed since the consumer applied to a debt counsellor for debt review. See further Scholtz *et al* para 11.3.3.3.

⁶² For a discussion of these powers, see Scholtz *et al* para 12.9.

5.2 Where the acknowledgement of debt is entered into with a juristic person

As indicated above, where an acknowledgement of debt is entered into as a small or intermediate section 8(4)(f) credit transaction with a small juristic person, the NCA will have limited application to the acknowledgement of debt as set out in section 6 of the Act. This will *inter alia* entail the following:

- (a) The provisions relating to registration as a credit provider in accordance with section 40, as discussed above, will apply.
- (b) The provisions relating to unlawful agreements, provisions and supplementary agreements as set out in sections 89, 90 and 91 will apply with the exception of section 89(2)(b)⁶³ and section 90(2)(o).⁶⁴
- (c) The provisions of the Act relating to consumer rights will apply, as indicated above. This includes the provisions relating to confidentiality, personal information and consumer credit records.
- (d) As indicated above, the provisions of Part C of Chapter 6 relating to pre-enforcement requirements contained in section 129 will apply as well as the provisions regarding debt procedures in a court set out in section 130.
- (e) The courts will have extended powers, as indicated above, in enforcement proceedings that are afforded by section 130(4) of the Act.

Although the creditor will thus have a lighter compliance obligation than in the case of a natural person as he does not have to observe the provisions of the Act relating to interest, fees and other charges and will not have to deal with issues of reckless credit and over-indebtedness, as these are not to the avail of the juristic person debtor, it is clear that the creditor will still be saddled with various issues he would not have had to deal with if the acknowledgement did not constitute a credit agreement governed by the NCA.

6 INFLUENCE OF THE CAUSE OF ACTION THAT GAVE RISE TO THE ACKNOWLEDGEMENT OF DEBT

The cause of action in relation to which the acknowledgement of debt was entered into may be based either on contract or delict. Where it is based on a contract or agreement which constitutes a credit agreement, it is submitted that the insertion of a no-novation clause into an acknowledgement of debt will not serve to take the agreement subsequently concluded out of the ambit of the NCA. However, where the debt initially arose as a result of a delict, it is submitted that the insertion of a no-novation clause might have the effect of preserving the original cause of action, namely the delict, and thus cause the matter to fall outside the scope of the NCA.

7 CONCLUSION AND SUGGESTIONS

It is debatable whether the legislature considered the effect of the NCA on standard acknowledgements of debt during the drafting of the Act. It is, however, not unlikely that this was done given that the intention of the legislature was, *inter alia*, to extend protection by means of section 8(4)(f) to a consumer who

⁶³ S 89(2)(b) provides that a credit agreement entered into as a result of negative option marketing is unlawful.

⁶⁴ S 90(2)(o) provides that a provision which states or implies that the rate of interest is variable, except to the extent permitted by section 103(4), is unlawful.

enters into a credit agreement that falls outside the named credit transactions in section 8(4)(a) to (e).

Although one might argue that it could never be the intention of the parties, when arranging the repayment terms of an existing debt, to enter into a credit agreement, the fact of the matter remains that such acknowledgement of debt entails a deferral of payment (usually by means of monthly payments) in respect of which interest, fees and other charges are then levied, causing it to fall squarely into the catch-all net of section 8(4)(f). It should be noted that the definition of “credit” for purposes of the NCA does not necessarily entail that money should be advanced as a result of entering into the agreement, but explicitly includes deferral of money owed to a person, thus including money owed in respect of a pre-existing debt.

One may of course wonder whether the mere fact that a further “credit agreement” is extended to a consumer who is already unable to pay his debts and is clearly either over-indebted or likely to become so in future would not constitute reckless credit. The answer to this predicament would most probably be that the acknowledgement of debt should be viewed in substitution of the previously unpaid debt in respect of which it is entered, with the effect that such debt cannot be added into the over-indebtedness calculation together with the amount owing in terms of the acknowledgement of debt. Thus, if a credit assessment is done prior to entering into the credit agreement and the result thereof is that the debtor will be able to afford the repayment of the debt in accordance with the terms of the acknowledgement of debt, the fact that he was previously unable to settle that debt on its original payment terms will not cause the acknowledgement of debt to constitute reckless credit.

The practical reality is most probably that many creditors who enter into standard acknowledgements of debt with consumer debtors on a regular basis are, due to the nature of their core business, already registered as credit providers, such as banks who provide credit in terms of a variety of credit agreements. To them the further implications of treating an acknowledgement of debt as a section 8(4)(f) credit agreement, such as the compulsory prior assessment required in terms of section 81, the limitation of interest and the pre-enforcement and enforcement requirements might be less cumbersome given that they will already have structures in place to deal with these aspects. However, should the creditor for instance be a natural person who is not actually in the business of extending credit and who advanced a large amount of money in excess of R500 000 to another natural person on terms that are at arm’s length and the parties enter into an acknowledgement of debt for repayment of the loan amount, the implications for the creditor may be severe as he would be required to register as a credit provider and will have to comply with the applicable provisions set out in paragraph 5 above.⁶⁵

In an attempt to avoid the onerous application of the NCA to a specific standard acknowledgement of debt relating to repayment of an existing debt, it is submitted that the judgment in *Voltex (Pty) Ltd v Chenleza CC*⁶⁶ might probably

⁶⁵ It is conceded that much of the problems that this specific creditor would encounter would be as a result of the inequitable operation of the credit-provider registration requirements in s 40(1)(a) and (b) of the NCA in respect of creditors of the nature mentioned in this example. A detailed discussion of this section is beyond the scope of this article.

⁶⁶ 2010 5 SA 259 (KZP).

offer a solution, as the court held: “No charge, fee or interest was payable to the plaintiff in terms of the agreement, save the interest which was payable as damages in consequence of the breach of contract. Such interest was not fixed or determined by the agreement *but by operation of law*.”⁶⁷

Clearly, if the acknowledgement of debt only caters for a deferral without requiring payment of interest, fees and other charges, it will not comply with the requirements of section 8(4)(f) and will consequently be able to escape the application of the Act. This would, however, not be an economically feasible option for most creditors who have already been forced to wait for repayment of debts owing to them. It would consequently seem that such creditors would at least be able to recover *mora* interest at the rate of 15,5% as indicated in the Prescribed Rate of Interest Act⁶⁸ as damages in consequence of breach of contract without risking the NCA becoming applicable to the acknowledgement of debt.

Where the acknowledgement of debt is for the repayment of a debt based on delict, there is also, as pointed out above, the option of inserting a no-novation clause into the acknowledgment in order to preserve the delict as underlying *causa* and so avoid the application of the Act to the acknowledgement.

It is indeed true that a creditor to whom an amount is owing by a debtor can choose whether to enter into an acknowledgement of debt or not. If he does not wish to do so, he can institute action against the debtor and if he obtains judgment, he can then proceed with execution against the debtor’s assets. The latter choice is of course only viable where the debtor indeed has assets that would make it worthwhile to opt for litigation and execution. There are, however, a large number of debtors against whose assets execution is not an economically-viable option or who do not possess assets that can be executed against. Given the protection of a person’s right of access to adequate housing contained in section 26 of the Constitution of the Republic of South Africa, 1996, and the strict enforcement of this right by the Constitutional and other courts,⁶⁹ the prospects of executing against immovable property which is the debtor’s home have also become more restricted. It has been held that one of the factors that a court might consider when deciding whether to authorise a warrant of execution against the debtor’s home is whether the debt in regard to which the warrant is sought could have been satisfied by other reasonable means.⁷⁰ Entering into an acknowledgement of debt with the debtor would of course in many instances constitute such reasonable means, thus creating a catch 22-situation for many creditors effectively leaving them with no other option than to enter into an acknowledgement of debt with the debtor. It is, however, submitted that the creditor’s unenviable situation might be alleviated if at least it is possible to avoid the application of the NCA and its onerous protective provisions to acknowledgements of debt for repayment of existing debt by debtors who have already been shown their fair share of leniency.

67 Para 39; my emphasis.

68 55 of 1975.

69 *Jaftha v Schoeman; Van Rooyen v Stolz* 2005 2 SA 140 (CC); *Nedbank v Mortinson* 2005 6 SA 462 (W); *Standard Bank of South Africa Ltd v Saunderson* 2006 2 SA 264 (SCA); *Gundwana v Steko Development CC* 2011 3 SA 408 (CC) and *Firstrand Bank Ltd v Folscher* 2011 4 SA 314 (GNP).

70 See *Jaftha* paras 56–60.