Northern Ireland Legal Quarterly

Volume 70 Number 2

Summer 2019

Chief Editor

Dr Mark Flear

Queen's University Belfast

School of Law

NILQ

Contents

ISSN 2514-4936 (online) 0029-3105 (print)

NILQ 70(2): 181–201

A millstone around the neck? Stereotypes about wives and myths about divorce

SHARON THOMPSON

School of Law and Politics, Cardiff University[1]

Abstract

This article examines the effect of and connection between stereotypes about wives and myths about financial provision on divorce. It is based on an assessment of press reports on Mills v Mills *which have fuelled calls for reform of the law of financial provision on divorce, most notably the Divorce (Financial Provision) Bill 2017–2019. It argues that gendered stereotypes about wives (such as 'alimony drone', 'gold-digger' and 'meal ticket for life') inhibit shifts towards substantive equality for women. These gendered stereotypes have not only framed the debate, creating a non-existent problem for reform proposals to solve, but have also affected the way in which individuals form family law agreements, which is important given the current policy emphasis on private ordering. The article concludes by proposing that reflexive engagement with stereotypes about wives should be an important part of judicial reasoning.*

Keywords: financial provision on divorce; *Mills v Mills*; gold-digger; alimony drone; meal ticket for life; judicial notice; nuptial agreements; financial settlements

Introduction

In 1983, Juanita Frances (founder and then President of the Married Women's Association)[2] wrote: 'It is said the media's campaign is a backlash against the modicum of equality women have so far achieved.'[3] She was referring to the Justice in Divorce campaign in the 1980s that contributed to the introduction of the 'clean-break' principle[4] which facilitated the ending of financial ties between divorcing spouses (and in particular the payment of spousal support).[5] This Justice in Divorce campaign was not about the

1 Earlier versions of this paper were presented at the Universities of Durham, Warwick, Bristol and NUI Galway between 2017–2018. I would like to thank the attendees at these talks for their valuable comments, and to Dr Andy Hayward, Dr Illan Wall, Dr Katie Cruz and Dr John Danaher for inviting me to present this research on those occasions. I am also indebted to Professor Russell Sandberg for his detailed feedback on this article, and I am very grateful to the anonymous reviewers for their insightful, helpful and comprehensive comments.

2 The Married Women's Association was formed in 1938 and disbanded in 1988. It campaigned for marriage to be recognised as an equal financial partnership in law: Sharon Thompson, 'Married Women's Property Act 1964' in Erika Rackley and Rosemary Auchmuty (eds), *Women's Legal Landmarks* (Hart 2018) ch 36.

3 Letter from Juanita Frances (25 January 1983), London School of Economics, The Women's Library, Records of the Married Women's Association, 5MWA/5.

4 Matrimonial and Family Proceedings Act 1984, amending the Matrimonial Causes Act 1973.

5 Gillian Douglas, *Obligation and Commitment in Family Law* (Hart 2018) 120.

severe impact separation could have on wives; it was about the alleged injustice of divorce for men.[6] The 'lively'[7] campaign received extensive media coverage, and Frances was right to be concerned about the consequences of this when she said:

> The controversy has vilified women and branded them with a 'meal ticket for life' syndrome. Accusations in the press have made them out as 'alimony drones', and the public has responded with concern that there could not be smoke without fire. Both men and women have jumped into the fray, showing bias and prejudice against the true state of the economically depressed position of women and children in divorce.[8]

Frances incisively captures the harm caused by stereotypes about women on divorce. The media's portrayal of wives as unfairly leeching from their ex-husbands in the 1980s[9] influenced public opinion and undermined the spirit of the Matrimonial Property and Proceedings Act 1970 which recognised the property entitlements of spouses as a result of their unpaid work in the home. Gillian Douglas notes that the many complaints about wives receiving 'meal tickets for life' came from 'disgruntled husbands and their second wives' and that the Law Commission focused on their hardship.[10] However, contemporary research findings suggested that the amount of maintenance transferred to former wives was significantly lower than media reports suggested.[11]

It is disappointing, yet predictable, given the formal equality and neoliberal influences discussed later in this article, that more than 35 years later, media coverage of Baroness Deech's Divorce (Financial Provision) Bill 2017–2019 make Frances' words resonate loudly. Once again, there is a media backlash against the move towards equality as articulated in 2000 by the House of Lords in *White v White*.[12] The power of this backlash is perpetuated by Baroness Deech who has argued her reform is urgent because:

> The wife who is least likely ever to have put her hand in cold water during the marriage is the one most likely to walk off with millions, regardless of her contributions or conduct. Hence we find that London is the divorce capital of the world for the wealthy, and the phrases 'gold digger' or 'alimony drone' have been coined.[13]

Drawing on my earlier research on the historical context of the gold-digger,[14] this article will take a closer look at this media backlash. For the first time, this article explores the use of stereotypes about wives in press reports concerning a judicial decision and how these stereotypes have fuelled calls for reform. In part one, the most prevalent stereotypes

6 Stephen Cretney, *Family Law in the Twentieth Century: A History* (Oxford University Press 2003) 139.

7 Douglas (n 5) 121.

8 Frances (n 3).

9 See Andrew Gilbert, *British Conservatism and the Legal Regulation of Intimate Relationships* (Hart 2018) ch 3.

10 Douglas (n 5) 121. Law Commission, *The Financial Consequences of Divorce: The Basic Policy, A Discussion Paper* (Law Com No 103, 1980). For a contemporary summary of the impact of the clean-break principle's introduction in 1984, see Pamela Symes, 'Indissolubility and the Clean Break' (1985) 48(1) Modern Law Review 44.

11 As Miles and Hitchings have pointed out, while advocates of reform complained of former wives unfairly and parasitically exhausting their ex-husband's resources, Maclean and Eekelaar's research showed that 'the amounts of maintenance transferred were very low, and constituted a much lower percentage of payors' household income than of recipients': Joanna Miles and Emma Hitchings, 'Financial Remedy Outcomes on Divorce in England and Wales: Not a "meal ticket for life"' (2018) 31(2) Australian Journal of Family Law 43–80, citing John Eekelaar and Mavis Maclean, *Maintenance after Divorce* (Clarendon Press 1986).

12 [2000] UKHL 54.

13 HL Deb 27 June 2014, vol 754, col 1491.

14 Sharon Thompson, 'In Defence of the Gold-digger' (2016) 6(6) Oñati Socio-Legal Series 1225–1248.

about wives in the media are introduced and contextualised. Following this, part two takes a closer look at how such stereotypes are utilised. It will do this by studying six national press reports published online over the course of one week[15] following the Court of Appeal decision in the financial remedies case *Mills v Mills*.[16] These press reports are evaluated to ascertain how stereotypes about wives are communicated to readers, and so the focus is on coverage of *Mills* rather than an analysis of the case itself. Though *Mills* has now been decided in the Supreme Court,[17] the Court of Appeal coverage is studied instead to assess initial reporting of the facts of the case.[18]

Press coverage of financial provision on divorce arguably plays an important role in shaping public perceptions,[19] and so, in part three, the effect of stereotypes prevalent in the press reports from part two is further explored. It will be argued that the unsubstantiated fear of terms like 'alimony drone', 'gold-digger' and 'meal ticket for life'[20] have led to continued calls for sweeping reform (including binding prenuptial agreements)[21] and claims that England and Wales are undesirable places for wealthy men to get divorced.[22]

In addition to highlighting the ramifications of stereotypes about women and the subsequent myths promulgated on divorce, along with their practical and damaging impacts, in part four this article also confronts the question of what can be done to counter them. The aim of this analysis is not only to expose the empirical evidence undermining claims of gold-digging in financial provision cases. Rather, it is to follow the work of Rosemary Hunter, Angela Melville[23] and others,[24] who have sought to challenge family law 'post-truths'. As a result, this article contends that negative stereotypes about women are dominant because those who talk about tropes such as gold-digging believe incorrectly that it is a widespread social problem. These perceptions are increasingly effective in family law, as the withdrawal of legal aid[25] means couples increasingly must reach their own financial settlements on divorce.[26] When these settlements depend on

15 6–12 February 2017.

16 [2017] EWCA Civ 129.

17 *Mills v Mills* [2018] UKSC 38.

18 Much of the reporting on the appeal to the Supreme Court and the Supreme Court decision repeats the information published in the initial reporting of the case.

19 This is also argued in the context of the criminalisation of paid homecare workers in Lydia Hayes, 'Sex, Class and CCTV: The Covert Surveillance of Paid Homecare Workers' in Lisa Adkins and Maryanne Dever (eds), *The Post-Fordist Sexual Contract: Working and Living in Contingency* (Palgrave 2015) 171, 183.

20 In particular, see Miles and Hitchings (n 11), which outlines findings from a range of data sets showing that the meal-ticket-for-life concern is unfounded.

21 Divorce (Financial Provision) Bill 2017–2019; Frances Gibb, 'Family Matters: Urgent Call for New Divorce Laws as Judges Demand Overhaul of "Corrosive" System' *The Times* (London, 19 November 2017) <https://www.thetimes.co.uk/edition/news/urgent-call-for-new-divorce-laws-as-judges-demand-overhaul-of-corrosive-system-8k0ncg7gt>.

22 See, for instance, the second reading debate of the Divorce (Financial Provision) Bill 2017–2019, where Baroness Deech reiterated the now commonly used phrase that London is the 'divorce capital of the world': HL Deb 11 May 2018, vol 791, col 376.

23 Rosemary Hunter and Angela Melville, '"As Everybody Knows": Countering Myths of Gender Bias in Family Law' (2001) 10(1) Griffith Law Review 124.

24 Maebh Harding and Annika Newnham, 'Initial Research Findings: The Typical Levels of Parental Involvement where Post-separation Parenting is Resolved by Court Order' (2014) 44(5) Family Law 672; Julie Doughty, Lucy Reed and Paul Magrath, *Transparency in the Family Courts: Publicity and Privacy in Practice* (Bloomsbury 2018); Miles and Hitchings (n 11).

25 Pursuant to the Legal Aid, Sentencing and Punishment of Offenders Act 2012.

26 See Emma Hitchings, 'Official, Operative and Outsider Justice: The Ties that (May not) Bind in Family Financial Disputes' (2017) 29(4) Child and Family Law Quarterly 359.

parties' respective negotiating strengths, the non-moneyed spouse's bargaining position is weakened by the perception that she will be an undeserving gold-digger or parasitic alimony drone if she asks for more money than her spouse is willing to part with. If the influence of stereotypes about wives is therefore not fully recognised, our understanding of the context in which settlements are reached on divorce is incomplete. It is imperative that the effect of gendered stereotypes on these settlements is recognised because, as contract theorists like Eric Posner assert, the reality in which parties make agreements is complex and 'powerful social norms may play a greater role in regulating contracts than the law does'.[27] Crucially, inaccurate stereotypes are harmful as they mis-frame and diminish the real issues of economic disadvantage on divorce experienced by those with caring responsibilities. Reflexivity is therefore essential in judicial reasoning, to acknowledge this broader context and to tackle tropes head on in judgment-making.

1 Stereotypes about wives

'Alimony drone' and 'gold-digger' are gendered phrases in that they are applied almost exclusively to women. Even the definitions of 'gold-digger' and 'alimony drone' in the *Oxford English Dictionary* (OED) refer explicitly to women and not men.[28] The OED defines 'alimony drone' as: 'A woman who lives on alimony payments and makes no attempt to support herself.'[29] Similarly, 'gold-digger' is defined in the OED as: 'A woman who forms relationships with men purely to obtain money or gifts from them.'[30] These definitions tell us that, first, women to whom these labels are attached are either deliberately economically dependent or are concerned solely with material gain and, secondly, that men are not alimony drones or gold-diggers. As a result, when these terms are employed in the media, they are used to vilify women, not men. Ultimately, and as the following analysis illustrates, these terms lead to women's intentions being questioned when entering relationships like marriage and when negotiating financial settlements on divorce. Even worse, the phrases 'gold-digger' and 'alimony drone' can be used arbitrarily regardless of women's intentions when marrying or divorcing. This is evidenced in Terry Arendell's research on the enactment of gender in interviews with divorced men. She found that 'most … interviewees believed, and insisted repeatedly, that women unjustly "soak" or "bleed" men dry economically on divorce'.[31]

This is also inextricably linked to the common perception that women are routinely granted a 'meal ticket for life' by the court on divorce. The 'meal ticket' phrase is generally associated with 'joint lives' orders, whereby spousal maintenance lasts indefinitely until

27 Eric Posner, 'A Theory of Contract under Conditions of Radical Judicial Error' (2000) 94 Northwestern University Law Review 749.

28 The term 'fortune hunter' has historically been applied to men, but the dictionary definition is gender neutral: see <https://en.oxforddictionaries.com/definition/fortune_hunter>. This is interesting given that nineteenth-century literature frequently casts the fortune hunter as a man seeking to marry a woman of wealthy status (see, for example, Mr Wickham in Jane Austen, *Pride and Prejudice* (T Egerton 1813)) when 'gold-diggers' in literature are almost (if not) always women (see, for example, Holly Golightly in Truman Capote, *Breakfast at Tiffany's* (The Folio Society 2013), a character often labelled as being a gold-digger). While the term 'gold-digger' continues to be part of our vernacular, 'fortune hunter' is not.

29 OED <https://en.oxforddictionaries.com/definition/alimony_drone>.

30 OED <https://en.oxforddictionaries.com/definition/gold-digger>.

31 Terry Arendell, 'Reflections on the Researcher–Researched Relationship: A Woman Interviewing Men' (1997) 20(3) Qualitative Sociology 341.

the court orders otherwise, the recipient re-partners[32] or one of the parties dies.[33] Quantitative and qualitative research[34] by Joanna Miles and Emma Hitchings shows that such orders are rare and that, more broadly, spousal support is generally awarded to mothers caring for minor dependent children. Overall, they found that women continue to be economically disadvantaged following divorce.[35] Their research shows that, for instance, *The Times*' reporting on the 'propensity of the family courts to make spousal maintenance awards for life'[36] is wrong. Instead, the opposite is true – all their data indicated a 'clean break culture'. From a data set of 399 court files, joint lives orders were made in just 5.5 per cent of cases. Similarly, in Hilary Woodward's survey of 369 court cases, only 2 per cent were joint lives orders for more than a nominal amount.[37] This indicates that whilst gold-digger, alimony drone and meal-ticket-for-life tropes are made credible by media reporting, such reports do not reflect the reality for women on divorce.

Believing in these stereotypes has consequences. As Sandra Fredman put it:

> Stereotypical images of women have throughout the ages been used to justify detrimental treatment. Because women are classified as different in the relevant respects, it appears justifiable to subject them to detrimental treatment.[38]

Ingrained by long-standing biases, the stereotypes of alimony drones, gold-diggers and meal tickets for life have a personal cost for women. They are contributing to a shift in how women's economic entitlement is perceived on divorce and assist in justifying detrimental treatment of women. Perhaps most significantly, such tropes can affect the actual resolution of financial disputes on divorce. Hunter et al have shown that disputes are often affected by the emotional and moral norms of the parties.[39] Their empirical findings drawn from interviews with parties with experience of dispute resolution[40] indicated that for women these included 'feelings of guilt, pragmatism, sacrifice or self-preservation, concerns about compensation or a desire for reasonableness, which were rarely put forward by men'. In contrast, Hunter et al found that men expressed the norm of formal equality twice as often as women. This context of formal equality masks dependency and is tied to the trope of the independent woman, celebrates independence, and encourages women to react in a visceral way to the accusation of the gold-digger, by either seeking to prove they don't need their partner's money, or by vilifying women who do. Yet, as Joan Tronto argues, assuming the existence of autonomous actors (as most democratic political theories do) does not explain 'how individuals can balance autonomy

32 The maintenance will commonly terminate following long-term cohabitation and will always terminate upon remarriage.

33 See s 28(1) of the Matrimonial Causes Act 1973.

34 Drawn from interviews with practitioners in England in 2012–2013, two focus group discussions conducted in 2016 with 14 first instance judges and a court file survey of *c* 400 cases collected in 2012 from cases concluded between 2010–2012.

35 Miles and Hitchings (n 11). This finding is consistent with empirical research since the 1980s, discussed later in this article.

36 Michael Gouriet, 'Meal Ticket Appeal: Court Endorses a Clean Break' *The Times* (London, 26 April 2018) <https://www.thetimes.co.uk/article/meal-ticket-appeal-court-endorses-a-clean-break-nvm3gl20p>.

37 Hilary Woodward, '"Everyday" Financial Remedy Orders: Do They Achieve Fair Pension Provision on Divorce?' (2015) 27(2) Child and Family Law Quarterly 151, 164.

38 Sandra Fredman, *Women and the Law* (Clarendon Press 1998) 301.

39 Rosemary Hunter, Anne Barlow, Janet Smithson and Jan Ewing, 'Law, Discretion, Gender and Justice in Out-of-Court Financial Settlements' (2018) 31(2) Australian Journal of Family Law 189–203. See also Emma Hitchings, Joanna Miles and Hilary Woodward, *Assembling the Jigsaw Puzzle on Divorce* (University of Bristol 2013) for findings on how non-legal factors influence the process of dispute resolution.

40 This was part of a broader study by Hunter et al investigating out-of-court family dispute resolution processes and outcomes in England and Wales in both children's and financial matters between 1996 and 2014.

and dependency in their lives'.[41] Neither does it acknowledge the network of relationships and responsibilities that inherently impact decision-making surrounding the negotiation of financial settlements on divorce.[42] Therefore, formal equality and the assumption that everyone is equally autonomous 'reiterates well-worn patterns of discriminatory attitudes'.[43] As Emma Hitchings has aptly put it, formal equality's appearance of neutrality 'highlighted through the use of terms such as fairness, equality, settlement and autonomy' can mean that women are 'particularly vulnerable during financial settlement negotiations'.[44]

This is exacerbated by the dichotomy of 'gold-digger' and 'independent woman' often employed in media reports on divorce which contributes to the voices of the broad spectrum of women in between being silenced and lost. This includes an individual who becomes economically dependent on her partner because marriage *does* create interdependence, especially when there are children. She might make decisions for the welfare of the family that mean sacrificing her earning power.[45] And whilst she is not a gold-digger nor entirely self-sufficient, she is likely to be affected adversely by these powerful tropes.

As a result, this article investigates how stereotypes about wives have helped create a backlash against efforts to implement substantive equality and recognise care. This is significant because it leads to false perceptions about the law, which in turn seeps into public consciousness and could affect how parties negotiate financial settlements and future reform. To this end, I will first examine press coverage of wives and financial provision on divorce through a study of *Mills v Mills*.[46]

2 *Mills v Mills* in the news

My earlier work has shown how the phrase 'gold-digger' has been used as a weapon against women, which creates an image of wives as predatory and obscures the causes of economic dependency in the relationship.[47] To understand the wider cultural context of such images, and to help reveal how stereotypes enter the public domain, this section examines six national press reports[48] concerning the 2017 Court of Appeal financial provision case *Mills v Mills*.[49] The media reporting of wives in financial provision cases is an important part of this story. These reports illuminate media misunderstandings of the law of financial provision on divorce that are in turn likely to encourage public misunderstandings, showing how stereotypes about wives are linked to myths about the law. An awareness of this is important.[50] When decisions are made over financial settlements, nuptial agreements and the terms of such agreements, negotiations and the

41 Joan C Tronto, *Caring Democracy: Markets, Equality and Justice* (New York University Press 2013) 31.

42 Jonathan Herring, *Relational Autonomy and Family Law* (Springer 2014) 68. See also Sharon Thompson, *Prenuptial Agreements and the Presumption of Free Choice* (Hart 2015) ch 6.

43 Tronto (n 41) 31.

44 Hitchings (n 26) 361.

45 See Anna Heenan, 'Causal and Temporal Connections in Financial Remedy Cases: The Meaning of Marriage' (2018) 30(1) Child and Family Law Quarterly 75.

46 *Mills* (n 16).

47 Thompson (n 14).

48 This approach has been adopted in relation to press reporting on paid homecare workers: Hayes (n 19) 171. For more media reports on *Mills* (n 16) not included in this article, see: 'Mrs Millstone . . . ?' *The Transparency Project* (12 February 2017) <www.transparencyproject.org.uk/mrs-millstone>.

49 *Mills* (n 16).

50 Doughty et al (n 24).

exercise of power between the parties are affected by the gulf between public perceptions of the law[51] and how the law operates in practice.[52]

To emphasise the impact of media reporting on our understandings of events related to financial relief proceedings, I deliberately do not outline the facts of *Mills v Mills* before presenting media reports on this case so that these reports can be taken at face value. The sample of reports was put together by researching the 10 most-read news titles in the UK according to Ofcom.[53] From these titles, four sources were selected which covered *Mills* more than other titles,[54] and online articles published by these news titles were selected over a week-long period following the Court of Appeal decision (6–12 February 2017). The sample focuses on Court of Appeal coverage, as reports on the Supreme Court decision repeat much of the initial reports.

Headline 1: '*Divorced men doomed to life as a cash machine.*'[55]

This article presents the facts of this case as being simply that Mr Mills divorced his wife 15 years ago but has had to increase his payments to her because she had spent her divorce settlement. The phrase 'gold-diggers' is used in the article, and the author is highly critical of the law:

> It's often claimed that ex-wives can see their former husbands as cash machines, compounding this by stripping them of the human right of fatherhood. This not only flies in the face of equality but can also cause men to question their desire to live. In the past seven years suicides among men aged 45–49 have risen by 40%.

This report makes it clear that the outcome for the husband in *Mills v Mills* was in the author's view unfair and then uses this case to make a broader statement about the current law. The author calls for reform of the law of financial provision on divorce, connecting its injustice to male suicides. The headline and article misrepresent the joint lives order at issue in *Mills*, by explaining it as being 'an allowance she will enjoy for the rest of her life'. If Ms Mills re-partners or if her circumstances change, the order can be reviewed.

Headline 2: '*Court orders man to increase payments to wife who lost bulk of divorce settlement with poor financial decisions.*'[56]

The facts of *Mills v Mills* are ostensibly outlined in this report, whereby Ms Mills 'leaves the marriage with all, or almost all the liquid capital', subsequently loses it, and successfully argues that her periodical payments should be increased. In this article there is detail on Mr Mills' calls for reform of the law, seeking 'changes to the law to limit

51 For instance, there appears to be a common assumption that all assets are automatically split 50/50 on divorce under English law, as indicated on several family law practitioner websites (including: <https://harrogatefamilylaw.co.uk/is-everything-always-split-5050-in-divorce/>; www.portner.co.uk/news/divorce-myth-busting-the-couple-s-assets-are-always-divided-50-50).

52 The financial outcome on divorce is rarely 50/50 and instead depends on exercise of judicial discretion under s 23 of the Matrimonial Causes Act 1973 with a list of factors for the judge to consider under s 25 of the Act. For more detail on this, see Thompson (n 14) 1237.

53 Ofcom, *News Consumption in the UK: 2018*, 25 July 2018 <www.ofcom.org.uk/research-and-data/tv-radio-and-on-demand/news-media/news-consumption>. As reports were taken from online sources, the *Daily Mail* and *Mail on Sunday* were treated as being the same source because they were published on the same website.

54 For instance, by 1 January 2019, *Mills* had been reported 14 times in *The Times*, 9 times in the *Daily Mail*, 6 times in *The Telegraph* and 4 times in the *Evening Standard*. There were no online news reports in widely read news titles like *The Guardian*.

55 Martin Daubney, 'Divorced Men Doomed to Life as a Cash Machine' *The Times* (London, 12 February 2017) <www.thetimes.co.uk/article/divorced-men-doomed-to-life-as-a-cash-machine-lw526hj3n>.

56 'Court Orders Man to Increase Payments to Wife who Lost Bulk of Divorce Settlement with Poor Financial Decisions' *The Telegraph* (London, 6 February 2017) <www.telegraph.co.uk/news/2017/02/06/court-orders-man-increase-payments-wife-lost-bulk-divorce-settlement/>.

spouses to five years maximum maintenance'. Although the phrases 'alimony drone' and 'gold-digger' are not used in the article, there are several references to the wife as being financially dependent on Mr Mills without any effort to be self-sufficient. The sense of injustice for Mr Mills is clear, with his QC quoted as saying: 'It is wrong in principle and in law that the wife should continue to depend, and indeed seek to increase, her dependence on the husband.'

Headline 3: '*Business man's fury as judge orders him to support his ex-wife for life after she blew £230k payout on homes.*'[57]

This article is based on an interview with Mr Mills, which therefore does not present Ms Mills' side of the story. Although other reports implied that Ms Mills had successfully applied to the court for more money, it is revealed in this report that the case was in fact brought by Mr Mills. He sought to have her periodical payments (spousal maintenance) reduced, arguing their adult son is now independent and saying he regretted the original deal made at the time of their separation. Mr Mills' view of the Court of Appeal decision is that it is 'grossly unfair', and he calls for reform of the law:

> I don't think it's right that after divorce you should be tied together for ever. I don't think you can move on mentally or physically with that tie in place. The law is wrong in that regard … I feel like I am paying for her mismanagement of finances … I'm angry and frustrated the system allows this to happen. I don't really want this to be about me and her, I think there's a wider issue about the law. I've no need to put her down, I'd rather the effort was concentrated on making the law work for people.

Headline 4: "'*I'm paying for HER mistakes": Husband who was ordered to support his ex-wife for life 15 years AFTER they divorced calls for a time limit on maintenance after she blew her £230,000 payout on bad investment.*'[58]

This report reinforces the sentiments expressed by Mr Mills in the *Evening Standard* article. Even the language used in the headline of this article is provocative – 'I'm paying for HER mistakes' clearly elicits sympathy for the husband, anger at the wife and a sense of injustice about the law. We are again told that Mr Mills is 'furious' at the 'grossly unfair' decision reached by the court and the entire article focuses on Ms Mills 'failures' and 'poor financial decisions'.

At this point, taking these reports at face value would lead one to believe that the law does not encourage men and women to be financially independent after divorce. The law is presented as making lifetime awards that unfairly prejudice men and enable women to be alimony drones. And if the Court of Appeal in *Mills v Mills* is increasing maintenance as a result of the wife's own poor judgement after relationship breakdown, it could be argued that the law is encouraging gold-diggers, and that there does need to be reform as Mr Mills claims. These assumptions are understandable, but further investigation into this case reveals that these press reports have missed key details about the case.

57 David Churchill and Jonathan Prynn, 'Business Man's Fury as Judge Orders Him to Support his Ex-wife for Life after She Blew £230k Payout on Homes' *Evening Standard* (London, 8 February 2017) <www.standard.co.uk/news/london/businessman-s-fury-as-judge-orders-him-to-support-his-exwife-for-life-after-she-blew-230k-payout-on-a3461391.html>.

58 Alex Matthews and Steph Cockroft, "'I'm Paying for HER Mistakes": Husband who was Ordered to Support his Ex-wife for Life 15 Years AFTER They Divorced Calls for a Time Limit on Maintenance after She Blew her £230,000 Payout on Bad Investment' *Daily Mail* (London, 8 February 2017) <www.dailymail.co.uk/news/article-4204214/Court-order-makes-man-pay-ex-wife-s-mistakes.html>.

Headline 5: '*Trolls target woman after divorce payout.*'[59]

This report is one of the few examples where the reader gleans Ms Mills' perspective on the case. She clarifies a few facts that were not made clear in other reports:

> It should be noted that since the divorce 15 years ago, I have never returned to the court to increase my maintenance, despite my financial difficulty and bad health and low earnings … This was Mr Mills's application to the court to reduce or eliminate the maintenance despite the fact our son is still in full-time education and still living with me at home.

Ms Mills also comments on the other press reports on her financial provision case and their effect:

> There have been a number of articles published recently in respect of this appeal which give a very one-sided version of the facts in this case … [t]he tone and content of the reporting has caused me considerable distress, with internet hate mail.

Headline 6: "'*I'm NOT a gold digger": The ex-wife vilified as a cold, cunning chancer after winning a huge rise in maintenance payments 15 YEARS after her divorce blasts back and says it's all her ex's fault.*'[60]

This headline is representative of many of those studied, in that it adds to the news rhetoric of stereotypes about wives. As Faith Gordon notes, paying attention to headlines like this is important because it summarises the most 'poignant' pieces of information in a way that attracts the reader's attention and encourages the reader to process the text in a particular way.[61] This headline is a notable example of this. Though Ms Mills' perspective is reported here, the wording of the headline continues to reinforce the derogatory image created throughout other media reports: Ms Mills as 'a cold, cunning chancer' who 'blasts back' at her ex-husband. This does not invite the reader to be sympathetic about Ms Mills.

In the main text of the report, Ms Mills reveals that the coverage of her case in the press has caused her to be labelled as a gold-digger:

> I feel like I have been through a character assassination, both at the hands of my ex-husband and the public. They've said I'm a gold digger. It's not true. I'm nothing of the sort. The past two years of fighting my ex-husband have been incredibly stressful and have taken their toll on me physically and emotionally.

Although *The Telegraph* reported[62] that Ms Mills had received all the liquid capital on separation, in this article the reader learns that she did not perceive the original settlement as fair, and she ostensibly did not leave the marriage on an equal footing with her husband:

> At the time of our divorce my husband claimed he couldn't afford to buy a house for himself and was renting a room. He suggested I should use £200,000 to buy a small, terraced home for me and my son. Yet immediately after everything was

59 Frances Gibb and Jonathan Ames, 'Trolls Target Woman after Divorce Payout' *The Times* (London, 10 February 2017) <www.thetimes.co.uk/edition/news/trolls-target-woman-after-divorce-payout-ddvqv7qrn>.

60 Olga Craig, "'I'm NOT a gold digger": The Ex-wife Vilified as a Cold, Cunning Chancer after Winning a Huge Rise in Maintenance Payments 15 YEARS after her Divorce Blasts Back and Says It's All Her Ex's Fault' *Mail on Sunday* (London, 11 February 2017) <www.dailymail.co.uk/femail/article-4215578/Maria-hits-winning-increase-maintenance.html>.

61 Faith Gordon, *Children, Young People and the Press in a Transitioning Society: Representations, Reactions and Criminalisation* (Palgrave 2018) ch 4.

62 See n 56.

> settled he bought a £350,000 house for himself, his girlfriend and her daughter – and paid for that daughter's upbringing from the age of eight until she was 23.

Taken together with the previous report in *The Times*,[63] the reader learns that, although the Mills' son is an adult, he is not independent in the way Mr Mills stated in the *Evening Standard*.[64] Furthermore, although Ms Mills fell into financial difficulty as the press reported, she was not the party who applied to court. It was Mr Mills who brought the case to court because he wanted to terminate (or reduce) the amount of money he was paying Ms Mills, which was £1100 a month. In answer, Ms Mills asked the court to consider increasing her periodical payments as she was unable to meet her basic needs and ended up winning the appeal, having had to borrow money from friends and family to be able to respond to Mr Mills' application to court. Therefore, Ms Mills does not appear to fit within the dictionary definition of 'gold-digger'. Yet this is what she has been vilified as through press coverage like headlines 1 to 4, where each report's exposition of limited facts provides a skewed picture of events.[65]

A millstone around the neck?

Uniting virtually all of the press reports on *Mills* is the argument that this case highlighted a legal injustice, in that gold-digging or alimony-droning ex-wives are routinely awarded meal tickets for life. The solitary case of *Mills v Mills* is cited as authority for this, even though the work of Miles and Hitchings has shown this assertion to be unrepresentative.[66] Consequently, accusations in the press like those surrounding the *Mills* case suggest there must be 'no smoke without fire'[67] and that the prevalence of 'meal ticket' orders on divorce is of real concern. And so with striking parallels to the Justice in Divorce campaign of the 1980s, *The Times* subsequently launched its own Family Matters campaign in late 2017 calling for 'urgent' and 'sweeping' reform which would include an 'end to the so-called meal ticket for life maintenance awards'.[68] 'The victim of the law' under the Justice in Divorce campaign, Douglas notes, was 'the wronged husband, left with the millstone of the ex-wife around his neck'.[69] The same picture is painted by media reports now, even though the facts in the cases being reported continue to tell a different story.

On closer analysis of the *Mills v Mills* Court of Appeal judgment,[70] a more balanced picture can be gleaned than the one portrayed in press reports. Ms Mills was awarded an increase in periodical payments to meet her needs,[71] an increase which was subsequently

63 Gibb and Ames (n 59).

64 Churchill and Prynn (n 57).

65 There are other reports in tabloids such as *The Sun* (Laura Burnip, "'IT'S GROSSLY UNFAIR" Ex-Husband Forced by Judge to "Support Ex-wife for Life" after She Blew £230,000 Divorce Payout Slams "Unfair" System' *The Sun* (London, 8 February 2017) <www.thesun.co.uk/news/2816831/ex-husband-forced-to-support-wife-for-live-slams-system/>) and the *Daily Mirror* (Ann Stenhouse, 'Divorcee Who Blew Settlement Money Wins Higher Maintenance from Ex – and He Must Support Her for LIFE' *Daily Mirror* <www.mirror.co.uk/news/uk-news/divorcee-who-blew-settlement-money-9769361>) which fit the sample frame and repeat many of the aspersions cast on Ms Mills but do not add anything new to the analysis in this section, other than to further illustrate how stereotypes about wives are communicated to readers in mainstream news outlets.

66 Miles and Hitchings (n 11).

67 Frances (n 3).

68 Gibb (n 21).

69 Douglas (n 5) 124.

70 For a more detailed account of this case, see Heenan (n 45).

71 In 2002, the parties' financial claims were settled on divorce by a consent order, providing Ms Mills with a capital sum and periodical payments of £1100 pcm for joint lives. The case currently before the court was brought by Mr Mills in 2014, who sought to discharge or reduce the periodical payments being made to Ms Mills.

overturned by the Supreme Court in 2018.[72] The Court of Appeal found that, although Ms Mills had made unwise investments,[73] there had not been financial mismanagement,[74] and the court was clear that Ms Mills had not been 'profligate or wanton'.[75] There was nothing to suggest that Ms Mills was a gold-digger or alimony drone, nor did her unwise investments indicate mercenary intentions. Yet, in all six press reports assessed above, Ms Mills is either castigated for being a gold-digger, or (in the reports where she gives her own account) she is defending herself from the gold-digger accusation. This is because when the parties first separated a joint lives order was made (interpreted in the press as a meal ticket for life). This joint lives order was not affected by the Supreme Court's decision in this case; the court simply restored the level of maintenance to the level it had been before the Court of Appeal decision.[76] The vilification of Ms Mills for receiving this award does not provide space for discussion of the potential utility of a joint lives order in some cases. It may well be an appropriate response to enduring economic disadvantage following relationship breakdown, especially given Miles and Hitchings' findings that:

> . . . a clean break does not necessarily indicate that the economically weaker party can be fully self-supporting. She may only be able to avoid reliance on the ex-husband by relying instead on others, or she may have no obvious adequate alternative means of support at all.[77]

However, the press's panic about meal tickets for life suggests that joint lives orders are never justified. On the rare occasion that a joint lives order is made (or, in the *Mills* case, is not discharged), the recipient (if a woman) will be labelled a gold-digger or alimony drone regardless of her intentions or whether the recipient fits the dictionary definition of these terms.

This is not to say there are not women who fit the dictionary definition of gold-diggers and alimony drones, but equally, there might be men entering marriage with mercenary intentions too. However, these stereotypes, by their very definition, make us question women, not men. Irrespective of this, English law does not encourage gold-digging, but it *does* have the capacity to recognise care. Both spouses, especially after long marriages, should be able to leave the marriage independently and on an equal footing economically.[78] The court will not order a clean break if this will leave one party much worse off that the other.[79] Yet the 'clear "steer" in favour of the clean break' noted by Douglas[80] indicates that, in recent case law,[81] there is a firmly embedded expectation that wives will become self-sufficient following relationship breakdown.

The *real* gold-diggers, therefore, will be outwitted if they believe the myth that under English law the spouse will always get half of all assets on divorce or a meal ticket for life. If the court hears that the marriage has been short, and that little contribution has

72 *Mills* (n 17).

73 *Mills* (n 16) [12].

74 Ibid [13].

75 Ibid.

76 *Mills* (n 17).

77 Miles and Hitchings (n 11).

78 *Miller v Miller; McFarlane v McFarlane* [2006] UKHL 24.

79 A clean break is more difficult where there is less property and capital to divide up because there must be enough to meet the needs of the parties. See Emma Hitchings, 'The Impact of Recent Ancillary Relief Jurisprudence in the "Everyday" Ancillary Relief Case' (2010) 22(1) Child and Family Law Quarterly 93.

80 Douglas (n 5) 127.

81 *Matthews v Matthews* [2013] EWCA Civ 1874; *SS v NS (Spousal Maintenance)* [2014] EWHC 4183 (Fam); *L v L (Financial Remedies: Deferred Clean Break)* [2011] EWHC 2207 (Fam); *Wright v Wright* [2015] EWCA Civ 201.

been made to the marriage, the non-moneyed spouse will not be entitled to a large award.[82] Taking all of this into account, the media backlash indicated by press coverage of cases such as *Mills v Mills* has helped perpetuate a moral panic about gold-diggers and alimony drones and has propagated myths about the law of financial provision on divorce.[83] This is considered further in the following section, where the consequences of myths about financial provision on divorce are inspected more closely.

3 Myths about financial provision on divorce

The assertion in press reports on *Mills v Mills* – that the law encourages gold-diggers and alimony drones – has encouraged a moral panic that financial provision in divorce is in urgent need of reform so that husbands can protect their property. Without any robust academic proof that there is an issue with gold-digging or alimony-droning wives, women are routinely castigated as such. Even worse, these charges obscure the research showing women *in fact* do worse than men on relationship breakdown across all levels of wealth.[84] Stereotypes about wives cause a backlash against the equality so far achieved under law.

In the landmark case *White v White,* Lord Nicholls said that, when it comes to dividing assets on divorce, non-financial contributions to the welfare of the family such as reproductive labour, domestic labour and care for dependants must be valued as much as economic contributions.[85] Recognising the value of unpaid work in the home is important, Lord Nicholls said, because the division of labour in the marriage is gendered, and when someone's financial contributions are deemed more important by the law, the person responsible for contributing to the non-financial parts of family life is left at a disadvantage.[86]

Unfortunately, in spite of such judicial attempts to say that care matters and should be valued economically, non-financial contributions are consistently devalued in society.[87] Recent calls for divorce reform would undermine judicial efforts to raise consciousness about the more gendered aspects of both financial provision on divorce and prenuptial agreements.[88] Importantly, these calls for reform are not only from Mr Mills and other disgruntled husbands who are given prominent voices in press reports. As this section will show, those in positions of authority are also advocating reform based on the flawed assertion that the law of financial provision on divorce encourages gold-diggers and

82 This was reinforced in *Sharp v Sharp* [2017] EWCA Civ 408.

83 Other financial provision cases where the wife has been portrayed as a gold-digger or alimony drone include *Crossley v Crossley* [2007] EWCA Civ 1491 (Amanda Cable, 'How I Saw off Serial Divorcée Susan Sangster, by Husband Number 4' *Daily Mail* (London, 25 February 2008) <www.dailymail.co.uk/femail/article-517365/How-I-saw-serial-divorc-e-Susan-Sangster-husband-Number-4.html>); *Miller v Miller* [2006] UKHL 24 and *Charman v Charman* [2006] EWCA Civ 1791 (which were used as examples of a flawed system of financial provision on divorce by Ruth Deech: 'It's Time to Update our Divorce Laws' *The Guardian* (London, 15 September 2009) <www.theguardian.com/commentisfree/2009/sep/15/divorce-law-maintenance-update>). Though Heather Mills has been cast as the archetypal gold-digger following *McCartney v McCartney* [2008] EWHC 401 (Fam), there is some evidence of stereotypes surrounding this case being challenged in the liberal press, though these reports are much less prevalent (see Kira Cochrane, 'Why We Love to Hate Heather' *The Guardian* (London, 20 March 2008) <www.theguardian.com/uk/2008/mar/20/mccartneydivorce.pressandpublishing>.

84 Hayley Fisher and Hamish Low, 'Recovery from Divorce: Comparing High and Low-income Couples' (2016) 30 International Journal of Law, Policy and the Family 338.

85 *White* (n 12).

86 Ibid [24].

87 See Heenan (n 45).

88 See *Radmacher v Granatino* [2010] UKSC 42, at [175]–[176] (Lady Hale).

alimony drones, further embedding stereotypes about women and the myths these stereotypes promulgate.

If gold-digging is pathologised as a female trait,[89] the ramifications for reform of financial provision can be concerning. This becomes clear when one places oneself in the position of someone who has only read press reports on Mr Mills' framing of the outcome of his case as unjust. This individual then goes on to read the views of Baroness Deech, who is one of the authorities on family law in the House of Lords and is one of the 'legal grandees'[90] supporting *The Times*' call for an end to meal-ticket-for-life awards.[91] In press reports Baroness Deech directly refers to *Mills v Mills*. She believes the outcome was wrong in this case and the law needs to be reformed:

> Our judges are being very old fashioned I'm afraid. They are over-chivalrous and the way they were in the 19th century. People wonder why, 15 years after a marriage has ended, one person has to keep paying money to another.[92]

Here, Baroness Deech is confirming suspicions about the judiciary after reading one-sided accounts of Mr Mills' plight. Her words are persuasive in a society that vilifies gold-diggers and celebrates (at least in this context) 'independent women':

> If there is one thing that stops women getting back on their feet and being treated seriously and equally at work, it is the assumption throughout the legal system that once a woman is married she is somehow disabled and incapable ever of managing on her own for the rest of her life. It is a very serious impediment to equality.[93]

On this view, the law of financial provision on divorce is obstructing gender equality by sending a message that women are not as capable of independence as men are. Here, Baroness Deech's perspective is different from the press reports concerned with injustice to married men, but it is no less powerful. Baroness Deech's assertion is that the law is equally unjust to men *and* women because the law infantilises women, making it difficult for them to shake the 'housewife' image and be taken seriously as 'independent women'.

Baroness Deech's argument that the current law is *anti*-equality is ostensibly persuasive. She says that the disadvantages experienced by married women in the public sphere are not attributable to their husbands and that our law legitimises the attitudes of employers who discriminate against women because, in her view, they are aware that the law rewards gold-diggers and alimony drones. And so Baroness Deech has proposed reform which, she says, could help women to be more equal and independent in society.[94]

89 Press reports about gold-diggers and alimony drones are undoubtedly influential, but the power of these stereotypes is strengthened much more when those in positions of authority claim there is proof women do want rich men. One pertinent example of this in Canada is Justice Bauman's acceptance as evidence that women had evolved the tendency to be gold-diggers, based on the work of evolutionary psychologists: *In Reference re: Section 293 of the Criminal Code of Canada*, 2011 BCSC 1588, 555 (a case on the legality of polygamy).

90 Gibb (n 21).

91 Her views are supported by Baroness Fiona Shackleton, another lifelong peer and family law expert (see HL Deb 11 May 2018, vol 791).

92 Patrick Sawyer, 'Divorces are Skewed by Judges' Outdated Chivalry, says Female Peer Pushing for Cap on Payments' *The Telegraph* (London, 12 February 2017) <www.telegraph.co.uk/news/2017/02/12/divorces-skewed-judges-antiquated-chivalry-says-female-peer/>.

93 Ibid.

94 HL Deb 11 May 2018, vol 791, col 379.

The Divorce (Financial Provision) Bill 2017–2019 aims to reform the law of financial provision on divorce by doing three things.[95] First, it provides for an equal division of assets acquired during the marriage (but excluding assets acquired before the marriage, and property acquired through gift or inheritance). Secondly, the Bill seeks to curb periodical payments to no longer than five years (unless this would lead to 'serious financial hardship'). Finally, the Bill would make prenuptial and postnuptial agreements binding.[96] The Bill includes standard procedural safeguards such as requirements of independent advice, cool-off periods and disclosure but, importantly, the Bill does not include any scope to set these agreements aside if their effect would be to leave one person in a position of need (which the court currently has discretion to do).[97]

The combined effect of these clauses is to reduce the discretion afforded to the judiciary to ensure substantive equality.[98] As Baroness Deech's views are based on a liberal and formal version of equality, her Bill elides structural inequalities that discriminate against those made economically dependent through care. By citing independence, individual responsibility and choice as desirable conditions over and above those connected to the ethic of care, proposals are directed to the upper and middle classes and would detrimentally affect the non-moneyed spouse. Talk of gold-diggers, alimony drones and independent women therefore takes place in a context of what Tronto calls 'privileged irresponsibility'.[99] Men and women who are in conditions of privileged irresponsibility can take on public duties and leave behind private duties.[100] Even after separation, they can afford to pay someone else to clean their house and help care for their children. And so those who are not in such conditions of privileged irresponsibility stand to lose the most from the prevalence of stereotypes about wives and the myths about divorce such stereotypes promulgate, which fuel support for reform like that proposed by Baroness Deech.

4 Countering myths in the courtroom and beyond

So far, this article has argued that, in order to understand how stereotypes like 'gold-digger' and 'alimony drone' impact on women, it is necessary to consider the discourses within which these stereotypes are situated. But it is also necessary to confront the potentially damaging effect of these stereotypes and the question of what can be done to counter them in practice.

Baroness Deech frequently refers to the gold-digger when advocating for the passage of her Private Member's Bill. Indeed, when she first introduced the Bill in 2014 she said it was necessary to fortify the system of financial provision on divorce against exploitation by the 'gold-digger' who has 'deprived irrationally [a man] of everything he had worked for'.[101]

This quotation is powerful because it corroborates the stereotypes about wives in press reports and further entrenches myths about divorce because these words are spoken

95 At the time of writing, this Bill had progressed through the House of Lords and was awaiting its Second Reading in the House of Commons.

96 These agreements currently have decisive weight unless this would, in the prevailing circumstances, lead to unfairness: *Radmacher* (n 88).

97 For an overview of the status of prenuptial agreements in England and Wales, see Thompson (n 42) ch 1.

98 For a more detailed critique of this Bill, see Sharon Thompson and Russell Sandberg, 'Common Defects of the Divorce Bill and Arbitration and Mediation Services (Equality) Bill 2016–17' (2017) 47 Family Law 425.

99 Tronto (n 41) 60.

100 Ibid.

101 HL Deb 27 June 2014, vol 754, col 1491.

in Parliament by a peer with expertise in family law. By stepping into the debate with concrete reform proposals, Baroness Deech has fuelled suspicions that the law is undeservedly generous to wives and has ostensibly provided a solution to this problem with her Bill. However, research has consistently contradicted Baroness Deech's view, proving that, instead of focusing on the mythical prevalence of meal tickets for life, a real matter of concern should be women's economically precarious position on divorce.[102] Economic dependency within the family is based on gendered social values and is not simply encouraged by the system of financial provision in England and Wales, just as career sacrifices made by women are not, as Baroness Deech has said, simply a 'matter of choice'.[103] Indeed, Baroness Deech's argument that such sacrifices are simply down to the choice of the individual is another hallmark of the 'privileged irresponsibility' within her brand of formal equality. The freedom of one's choices depends on the conditions an individual is in.[104] Choice does not mean the same thing as equality, according to Tronto, because equal opportunity 'is a myth if there is no equality of care for children'.[105] As a result, while research shows there are more female breadwinners than ever before, this should not lead to assumptions as to the modern division of labour in the home, especially when the marriage has produced children.[106] As Cynthia Lee Starnes has put it:

> According to myth, in today's egalitarian, gender-neutral culture mothers and fathers co-parent, both working full-time in the paid economy and sharing equally in their leisure time the few family tasks that are really necessary . . . [yet] Even when a married mother works outside her home she likely serves as the primary caretaker, undertaking a disproportionately large share of household chores.[107]

Taking all of this into account, the assertions in Parliament[108] that the law encourages gold-diggers and alimony drones is damaging. In the short term, it could lead to reform that would be disastrous for those primarily responsible for care in the marriage. In the longer term, as these assertions have gone unchallenged in the House of Lords since 2014 and are still consistently reported in the press, their legitimacy has increased. It is therefore important that the negative consequences of these stereotypes are properly factored into questions surrounding legislative reform of financial provision on divorce. At the very least, it is important that the myths about divorce created by these stereotypes are challenged directly in future parliamentary debates.

Unfortunately, stereotypes about wives and corresponding myths about divorce are not only a problem for debates around legislative reform. The impact on public perception of the law is also important since many people now determine the financial consequences of divorce for themselves, either by signing a nuptial agreement or by reaching a settlement on divorce.

---

102 See Heenan (n 45) 88.

103 Ruth Deech, 'What's a Woman Worth?' (2009) 39(12) Family Law 1140.

104 Thompson (n 42) ch 6.

105 Tronto (n 41) 41.

106 Office for National Statistics, *Families and the Labour Market, England: 2018*, 3 October 2018 <www.ons.gov.uk/employmentandlabourmarket/peopleinwork/employmentandemployeetypes/articles/familiesandthelabourmarketengland/2018>.

107 Cynthia L Starnes, 'Mothers, Myths, and the Law of Divorce: One More Feminist Cast for Partnership' (2006) 13(1) William and Mary Journal of Women and Law203, 208.

108 Not just by Baroness Deech. For example, Lord Davies asserted '[p]eople have been deeply offended by some of the gold-digging – that is the word one must use – that has had a lot of publicity recently': HL Deb (2014) 754, cols 1496–1497.

As I have argued previously, this was evident in my own empirical research.[109] When researching prenuptial agreements, I interviewed attorneys in New York and found that often the non-moneyed spouse signs a prenup to prove she is not marrying for money, or the moneyed spouse insists on a prenup to be reassured he is not marrying a gold-digger. Women are more inclined to sign agreements out of fear of confirming the alimony-droning, gold-digging stereotypes. To use one example from my interviews, the attorney explained that he had a client that used these stereotypes to his advantage:

> She did not want the prenup. [My client] insisted and at the last minute he capitulated and said it was alright. It was almost like it was a test. And because she said she would do it, [he knew] 'ok you're not marrying me for my money'.[110]

Why, as in this example, would a woman sign an agreement that she did not want to sign? One reason is that by not signing, women are faced with the threat of confirming stereotypes about wives. Focusing only on the motivations of the non-moneyed spouse not only leads to a false impression of the prevalence of gold-diggers and alimony drones, but also has a detrimental impact on financial outcomes on divorce. Parties may not ask for more money out of a fear of confirming such stereotypes. And so moral panics about gold-diggers and alimony drones add an extra layer of coercion to financial settlements and nuptial agreements. The recent English High Court case *KA v MA*[111] provides an example of the effect such stereotypes can have. In this case, the wife was 'psychologically torn' between proceeding with the wedding and signing an agreement 'which might well, at some point in the future, operate to her significant financial detriment'.[112] Roberts J considered the wife's reasons for entering the prenup, finding that the wife 'took this course for what may have been the best of reasons . . . motivated principally by what she perceived to be in their son's best interests'.[113] She even suspected that the wife did not make much attempt to ask for more under the agreement[114] out of fear that the wedding would not go ahead.[115] 'I am not like that',[116] the wife said. Taking the circumstances of this case into account, it was clear what she meant by 'that'. By refusing to ask for more money under the terms of the prenup, she would show she was not a gold-digger or a future alimony drone. She would not be labelled the 'cold, cunning chancer',[117] as Ms Mills had been in the *Mail on Sunday*. And her wedding could go ahead.

These examples indicate that tropes like 'alimony drone', 'gold-digger' and 'meal ticket for life' are influencing individuals' decision-making. When media narratives place emphasis on these tropes, the intentions of the non-moneyed (not the moneyed) party are interrogated in a way that does not place the non-moneyed parties' financial demands, needs and choices on a par with those of the moneyed spouse. In practice this can lead to spouses agreeing to financial settlements that will ultimately affect them detrimentally. Furthermore, focusing on whether someone is a gold-digger, or on requiring her to prove to the world that she will never be economically dependent (and signing entitlement away accordingly), facilitates a shift in emphasis away from valuing care and other non-financial contributions. The impact of devaluing care on relationship breakdown is that structural

109 Thompson (n 42) ch 3.

110 Ibid 196.

111 *KA v MA (Prenuptial Agreement: Needs)* [2018] EWHC 499 (Fam).

112 Ibid [63].

113 Ibid [65].

114 The husband increased his proposal once after the wife suggested it was 'a bit mean': ibid [63].

115 Ibid.

116 Ibid (emphasis added).

117 Craig (n 60).

inequalities are reinforced and those who do make the non-financial contributions to the welfare of the family are disadvantaged.

Research shows how gendered this work still is. In November 2016, the Office for National Statistics reported that women across all wealth brackets take longer to recover after divorce than men.[118] The work of Fisher and Low also indicates that women take longer to recover economically on relationship breakdown at all levels of wealth.[119] And so limiting financial provision to care-giving spouses is neither a safe nor productive strategy for addressing structural inequalities between men and women. But the meal-ticket, gold-digger mentality does not appreciate this reality.

However, confronting the question of what can be done to counter these tropes is difficult. Explaining the law on financial provision to those who are certain it makes women treat men like cash machines[120] feels impossible in the current climate. Not only is the law in this area complex; the press reports assessed above indicate that the voices of experts are being drowned out by more forceful voices shouting about injustice when they see women receive substantial awards on divorce. There is a perception that capital earned by the moneyed spouse is his alone, rather than capital generated directly and indirectly by both spouses working outside and inside the home. This is connected to the view that women who are awarded a share in this capital on divorce must be gold-diggers or alimony drones.

In spite of these difficulties, it is important that something is done. The influence stereotypes about wives have must be acknowledged and condemned in the courtroom. In *Waggott v Waggott*,[121] the wife's counsel raised this very issue, and according to Moylan LJ, he was:

> . . . evidently concerned that recent public debate about how the courts determine a spouse's claim for maintenance might somehow intrude into the determination of this case. His particular focus was what he described as the unfair use of the expression 'meal ticket for life' which, he suggested, was often deployed without regard to a spouse's fair entitlement which might properly include long-term maintenance.[122]

Ms Waggott's counsel was concerned that his client would be portrayed as an alimony drone or gold-digger for seeking an award that would place her on equal financial footing with her husband on divorce.[123] In response, Moylan LJ noted that he understood the expression 'meal ticket for life' 'can be used as an unfair trope' but refused to comment further. While it is commended that Moylan LJ recognised the unfairness of this stereotype, his judgment in *Waggott* could have engaged further with the influence of stereotypes in financial provision cases. Similarly, in his leading judgment in the Supreme Court in *Mills*, Lord Wilson said it was 'misleading' and 'unattractive' for joint lives orders to be described as being meal tickets for life but did not assess the impact of this phrase further.[124]

---

118 Office for National Statistics, *Women Shoulder the Responsibility of 'Unpaid Work'* (10 November 2016) <www.ons.gov.uk/employmentandlabourmarket/peopleinwork/earningsandworkinghours/articles/womenshouldertheresponsibilityofunpaidwork/2016-11-10>.

119 Fisher and Low (n 84).

120 Daubney (n 55).

121 *Waggott v Waggott* [2018] EWCA Civ 727.

122 Ibid [156].

123 She had left employment in 2001 and (apart from a short period in 2002/2003) had not worked in paid employment since, while the husband's career went from strength to strength: ibid [17].

124 *Mills* (n 17) [25].

Reflexive engagement with stereotypes should be an important part of judicial reasoning. There is a gap between what the law is and what social perceptions of the law are, and accurate information on the law[125] and evidence of the detrimental impact of tropes like alimony drone, gold-digger and meal ticket for life are important. Much has been written about the conditions that lead to power inequalities in intimate financial agreements, and conditions that lead to inequality of bargaining power are crucial to understanding how people make decisions in relationships.[126] But individuals' social perceptions are important too. Judgment-making has the potential to be an important influence in practice.

Judicial opinion will not counteract media stereotypes by ignoring them, and so there must be direct engagement with the negative consequences of such stereotypes on women. Failing to do so arguably obstructs justice which, as Tronto notes, 'requires honest exploration of the social, economic and political institutions that constrain people's lives'.[127] The influence of the moral panic surrounding stereotypes about wives is part of this context, as it directs attention towards the perception that overly generous awards are made and away from the detrimental impact of divorce on women.[128] This not only mis-frames important concerns, it also potentially reinforces the prevalence of clean-break orders in practice, which in turn affects the economic recovery of women after divorce. Unfounded stereotypes should not be relied upon at the expense of those in need of provision. By reflexively engaging with misplaced concerns in the courtroom, there is greater scope for members of the judiciary to interrogate parties' capacity for economic self-sufficiency instead of assuming that this will be possible following separation.

One way of doing this could be to bring 'social reality to the courtroom'[129] by taking 'judicial notice' of the gendered impact that stereotypes about wives can have. This would mean that the impact of derogatory, women-centred terms like 'gold-digger' and 'alimony drone' would be explicitly acknowledged, so that a factual inquiry into parties' actual (not assumed)[130] capacities to become self-sufficient can be scrutinised. Therefore, when considering whether to make a clean-break order, the court would invariably pay attention to evidence-based research of women's experiences and of empirical research on the financial impact of divorce and its impoverishing effect on women and children, instead of moral panics that have been shown to be ill-founded.[131] The doctrine of judicial notice, whereby a judge 'assume[s] a position of authority to find a "social fact" that

125 See, for example, the work of *The Transparency Project* at <www.transparencyproject.org.uk/>.

126 Alison Diduck, 'Autonomy and Family Justice' (2016) 28 Child and Family Law Quarterly 133; Anne Barlow and Janet Smithson, 'Is Modern Marriage a Bargain? Exploring Perceptions of Pre-nuptial Agreements in England and Wales' (2012) 24 Child and Family Law Quarterly 304.

127 Tronto (n 41) 41.

128 Miles and Hitchings (n 11).

129 Claire L Heureux-Dubé, 'Re-examining the Doctrine of Judicial Notice in the Family Law Context' (1994) 26 Ottowa Law Review 551, 551.

130 Miles and Hitchings' study found that a clean break is frequently ordered without clear absence of the parties' abilities to become self-sufficient (n 11).

131 This is an area where there is a clear contrast in family law between money and children cases due to the case management process. In children cases, the Children and Family Court Advisory and Support Service will bring in some relevant empirical evidence on child development to an individual case. See Bryan Rodgers, Liz Trinder and Teresa Williams, *Towards a Family Justice Observatory to Improve the Generation and Application of Research* (Nuffield Foundation 2015).

requires no reference',[132] is controversial (particularly in US jurisprudence)[133] in that it is an exercise of power because it depends upon his or her extra-legal, 'social' knowledge.[134] Acknowledging these difficulties, however, Claire Heureux-Dubé argues that the family law context is particularly well-suited to the use of judicial notice, as it enables the judiciary to assess 'the true consequences from the relationship and its breakup . . . formulating a more accurate picture of the realistic needs of the parties, particularly when self-sufficiency . . . [is] at issue'.[135] Using judicial notice in this way allows the judge to challenge stereotypes and to better understand the context in which financial matters are negotiated between parties on divorce. A possible counter-argument to this is that explicitly referencing stereotypes in judgment-writing could lead to accusations of bias in favour of wives. Claims of neutrality and objectivity in judicial reasoning are often criticised in feminist scholarship,[136] for members of the judiciary are inevitably influenced by their own experiences and opinions. Yet failing to directly acknowledge such stereotypes is arguably neither neutral nor objective either, when there is evidence they are detrimentally affecting women on divorce.

Appropriating the tools of feminist judgment-writing could assist in this regard.[137] One notable aspect of feminist judgment-writing is the central influence of care, emphasising the importance of relationships and interconnectedness instead of autonomy and independence.[138] This shift in emphasis is important given the neoliberal mantra that 'autonomy is good for you', which has repeatedly been used to legitimise the relegation of family disputes to the private sphere. The zenith of this change has been the virtual obliteration of legal aid in family law cases since 2012, a reform which Jess Mant argues is rooted in neoliberal ideology in that individuals are supposed to have responsibility for their own circumstances, whatever the structural barriers they experience.[139] The narrative of self-sufficiency in tandem with the backlash against dependency created by media stereotypes and Baroness Deech's proposed reform has created a hostile environment for the recognition and valuation of care. But by employing the tools of feminist judgment-writing, the courtroom can become a 'site of resistance to persistent gendered norms'[140] and stereotypes affecting the negotiation of financial settlements and nuptial agreements. Furthermore, as Judith Masson[141] and Duncan Kennedy[142] have argued, judgment-making is deliberative and strategic. Judges can shape 'not just an individual decision but the law itself through the focus and construction of

132 Heather Douglas, Francesca Bartlett, Trish Luker and Rosemary Hunter, 'Reflections on Rewriting the Law' in Heather Douglas, Francesca Bartlett, Trish Luker and Rosemary Hunter (eds), *Australian Feminist Judgments: Righting and Rewriting Law* (Hart 2014) 26.

133 Heureux-Dubé (n 129) 553.

134 Douglas et al (n 132) 26.

135 Heureux-Dubé (n 129) 576.

136 Douglas et al (n 132) 24.

137 Rosemary Hunter, Clare McGlynn and Erika Rackley (eds), *Feminist Judgments* (Hart 2010).

138 Douglas et al (n 132) 27. I also argue that this approach could be important for contract law in Sharon Thompson, 'Feminist Relational Contract Theory: A New Model for Family Property Agreements' (2018) 45(4) Journal of Law and Society 617.

139 Jess Mant, 'Neoliberalism, Family Law and the Cost of Access to Justice' (2017) 39(2) Journal of Social Welfare and Family Law 246.

140 Douglas et al (n 132) 26.

141 Judith Masson, 'Disruptive Judgments' (2017) 29(4) Child and Family Law Quarterly 401.

142 Duncan Kennedy, 'Freedom and Constraint in Adjudication' (1986) 36 Journal of Legal Education 518.

their judgments, even when another outcome of the case appears certain', and the content of judgments may be used for deliberate effect.[143]

Conclusion

Baroness Deech has consistently emphasised the need for reform of financial provision on divorce because of the perception that current law infantilises women. 'Younger academics . . . are reinventing the wheel', she says, by portraying women as victims.[144] Respectfully, I disagree. Concern for those with caring commitments in marriage is not based on antiquated ideas as suggested by Baroness Deech; it is a live issue. In the last decade, neoliberal policies have systematically devalued care and privileged those who do not care.[145] While the traditional dichotomy of housewife and breadwinner may no longer be as representative, normative expectations of care have instead resulted in women performing dual roles of both paid (often part-time) work *and* unpaid care-giving.[146] Recognising this and listening to the voices of the women (and men) who want their non-financial contributions to be valued is not to treat caregivers as victims; it is to recognise that a focus on maximising and protecting personal wealth ignores the realities of economic dependency in family life. However, this is not the story told in press reports. Instead, as this article has shown, wives are portrayed as alimony drones and gold-diggers who are parasitic on wealthy men. The phrase 'meal ticket for life' is employed repeatedly in these reports too, even though research has proven quantitatively and empirically that lifetime maintenance is not only rare,[147] but that women across all levels of wealth are economically worse off on divorce than men.[148] The research also suggests that proliferation of such stereotypes has created barriers for women when negotiating financial settlements and nuptial agreements, widening the gender gap on divorce even more.

Stereotypes about wives contribute to misunderstandings as to how the 'everyday' divorce[149] affects the spouse who has undertaken most of the non-financial work in the relationship, who is still usually the wife. And when the language of 'gold-digger', 'alimony drone' or 'meal-ticket for life' enters the debate, women must work much harder to justify getting what they are legitimately entitled to on divorce. Moreover, these stereotypes contribute to the drowning out of other concerns, such as Lord Nicholls' compelling call for equality after divorce in 2000: 'whatever the division of labour chosen by the husband and wife, or forced upon them by circumstances, fairness requires that this should not prejudice or advantage either party'.[150] Unfortunately, this judicial support for equality has not addressed the ongoing unequal economic impact of divorce on men and women.[151]

143 Masson (n 141) 404.

144 HL Deb 11 May 2018, vol 791, col 402.

145 See Anne Barlow, Jan Ewing, Janet Smithson, and Rosemary Hunter, *Mapping Paths to Family Justice: Resolving Family Disputes in Neoliberal Times* (Palgrave 2017).

146 Anne Barlow, 'Configurations of Unpaid Caregiving within Current Legal Discourse in and around the Family' (2007) 58(3) Northern Ireland Legal Quarterly 251, 252.

147 Miles and Hitchings (n 11).

148 Fisher and Low (n 84).

149 Hitchings (n 79).

150 *White* (n 12) [24].

151 In fact, as Miles and Hitchings' interview findings show, it is the norm of the clean break, not the meal ticket that prevails in legal practice, and this has 'produced outcomes that [appear] to have left some wives badly under-protected': Miles and Hitchings (n 11).

When inequality is bound up in complex political, social, economic and historical structures, there is no clear solution to closing the gap between men and women on divorce. However, negative stereotypes about wives are exacerbating these inequalities, creating false problems and simplistic solutions. As this article has argued, the court has a role to play in tackling the barriers created by such stereotypes. The case study of *Mills v Mills* is a clear demonstration of how myths about gold-diggers and alimony drones are propagated. Perhaps most worryingly, these press reports contribute to the misinformed perception that the law encourages and indulges the behaviour of money-grabbing wives. More broadly, judicial reasoning in financial provision cases should no longer ignore the impact that stereotypes about wives can have. Simply acknowledging these stereotypes is not enough; as this article has argued, judicial reasoning should go further and reflexively engage with the practical impact such stereotypes can have.[152] For, when these stereotypes are not addressed in the courtroom or challenged in Parliament, their harm is indirectly legitimised.

152 By, for example, taking judicial notice of empirical research on the financial impact on divorce, as discussed in this article.

NILQ 70(2): 203–219

Deception, mistake and non-disclosure: challenging the current approach to protecting sexual autonomy

OMAR MADHLOOM*

University of Bristol Law School

Abstract

English criminal law appears reluctant to criminalise deceptive sexual behaviour. It currently does so only in circumstances where the defendant has actively lied to the complainant regarding a fact recognised by law as crucial to consent. This restrictive approach arguably fails in many cases to protect the complainant's sexual autonomy. The central argument presented in this article is that all forms of deception, including non-disclosure, a false promise and mistake as to a material fact, may distort the complainant's decision-making process and undermine her ability to make an informed choice. A material fact is one which plays a significant role in a person's decision to engage in sex. This article advocates that the law of rape should be widened to include mistake on the part of the complainant and non-disclosure by the defendant.

Keywords: autonomy; consent; deception; lying; rape; sexual offences; undercover police

Introduction

In this article, I will argue that a complainant (C) should be deemed not to have consented to sexual activity in cases where any ostensible consent arises from deception[1] or misunderstanding, such as C's mistake, relating to a material fact. A material fact is one which plays a significant role in C's decision to permit or engage in sexual activity;[2] and it may be material to her, whether it would be material to someone else. A commitment to protecting sexual autonomy entails recognising that individuals are worthy of respect and, therefore, owed a duty not to have their sexual autonomy violated.[3]

Despite the self-evident value of sexual autonomy, the philosophical boundaries of acceptable behaviour have proven difficult to determine. Many definitions are reliant upon notions of 'consent', yet it is widely recognised that this term is fraught with

* My thanks to Irene Antonopoulos (De Montfort), Estella Baker (De Montfort), Gavin Dingwall (De Montfort), Michael Hirst (De Montfort), Susan Leahy (Limerick) and Ronnie Mackay (De Montfort) for their comments on earlier drafts of this article; to Lois Bibbings (Bristol), Peter Dunne (Bristol) and Vanessa Munro (Warwick) for their helpful discussions, and to the anonymous reviewers at NILQ. All errors remain my own.

1 For present purposes, 'deception' is synonymous with fraud, non-disclosure, false promise, mistake and misrepresentation.

2 David Archard, *Sexual Consent* (Perseus 1997) 46; Jonathan Herring, 'Mistaken Sex' [2005] Criminal Law Review 511, 518.

3 Home Office, *Setting the Boundaries: Reforming the Law on Sex Offences* vol 1 (Home Office 2000) paras 0.3–0.4.

conceptual ambiguity.[4] The current law of sexual offences only permits narrow forms of deception to negate consent; the defendant (D) must have 'actively' or 'deliberately' deceived C.[5] Section 76 of the Sexual Offences Act 2003 (the Act), allows for deception to negate consent in only two circumstances. Consent will be deemed to be absent where D 'intentionally deceived' C as to the nature or purpose of the relevant act,[6] or where C 'was intentionally induced' to consent by D impersonating a person known personally to her.[7] Deceptions falling outside these two categories may not negate consent. It is doubtful whether broader notions of deception, such as taking advantage of C's mistake or non-disclosure of a material fact, will be allowed to vitiate consent under s 74 of the Act. A person consents under s 74 'if he agrees by choice and has the freedom and capacity to make that choice'. Although courts use the term 'active deception', 'lying' is the more accurate term, as it involves making an assertion which D knows is false. Moreover, in *R v B*,[8] the Court of Appeal refused to hold that deception was analogous to non-disclosure.

It will be argued that, although lying and deceiving are philosophically distinct concepts, maintaining this distinction in the criminal law results in a failure to adequately protect sexual autonomy. Both lying and deceiving result in manipulating an individual's sexual choice by limiting the options available and preventing her from acting according to her own standards. The central argument in this paper is that the law of sexual offences should allow for all lies, deceptions, mistakes and non-disclosure relating to a material fact[9] to negate consent. Since the Act, undue rigidity has been applied by the courts in terms of the categories which negate consent. This suggests that the *R v Olugboja*[10] approach is preferable because it permitted the jury to decide in each case whether consent was present. This article will chart the unjustifiable approach to protecting autonomy.

The first section of this article addresses the relationship between sexual autonomy and non-consensual sexual offences, specifically the offence of rape, to show that it only permits narrow forms of deception to negate consent. The distinction between lying and deceiving will be examined to show that both concepts, while philosophically distinct, can violate sexual autonomy. It will then be argued that English law should allow broader notions of deception, such as mistake and non-disclosure, to vitiate consent. The article will conclude by arguing that the proposal can be accommodated within the existing law of sexual offences.

The current law of sexual offences

The origins of the Act lie in the recommendations of a Home Office review completed in July 2000. The focus of the review's recommendations was on personal autonomy, the prevention of sexual abuse or exploitation and the removal of discrimination in sex offences law.[11] Despite emphasising the importance of sexual autonomy, the review did not define this concept. This omission could be due to the fact that autonomy has been

4 Jennifer Temkin and Andrew Ashworth, 'The Sexual Offences Act 2003: (1) Rape, Sexual Assaults and the Problems of Consent' [2004] Criminal Law Review 328, 328; Victor Tadros, 'Rape without Consent' (2006) 26 Oxford Journal of Legal Studies 515, 521.

5 *Assange v Sweden* [2011] EWHC 2849 (Admin) [102]; *R v McNally (Justine)* [2013] 2 Cr App R 28 [21].

6 Sexual Offences Act, s 76(2)(a).

7 Ibid s 76(2)(b).

8 *R v B* [2006] EWCA Crim 2945, [21].

9 Archard (n 2) 46; Herring (n 2) 518.

10 *R v Olugboja* [1982] QB 320.

11 Home Office (n 3) paras 0.3–0.4.

described as a spacious word, capable of containing a variety of philosophical implications.[12] A common element exists amongst the various definitions; the emphasis on freedom of the individual. The review concluded that there should be a statutory definition of consent for the purposes of any non-consensual sexual offence. The definition of consent found in s 74 of the Act has been supplemented in certain circumstances by 'evidential' and 'conclusive' presumptions contained in ss 75 and 76. The review recommended that consent should be defined as 'free agreement'.[13] The Home Office review noted that the *Oxford English Dictionary* defines the verb 'to consent' as 'to acquiesce, or agree' and the noun 'consent' as 'voluntary agreement, compliance or permission'.[14] It recognised that these definitions cover a range of behaviour from whole-hearted enthusiastic agreement to reluctant acquiescence. However, in the context of sexual relationships, the core element is an agreement between two people to engage in sex.[15] By focusing on active deception, the current law fails to recognise the wide range of behaviours that are capable of negating consent.

The Home Office review recognised that:

> People have devised a complex set of messages to convey agreement and lack of it – agreement is not necessarily verbal, but it must be understood by both parties. Each must respect the right of the other to say 'no' – and mean it.[16]

The Law Commission had suggested that an apparent agreement should not count as consent unless it is a 'free and genuine agreement'.[17] It suggested that consent should be defined as 'subsisting, free and genuine agreement'.[18] This was rejected by the Home Office review as being too complex and introducing an unnecessary semi-contractual obligation complication into consent.[19] Instead, it recommended a definition based on 'free agreement' because of its simplicity and clarity. This, however, does not resolve the issue that 'freedom' and 'agreement' are ambiguous and complex concepts, which defy precise definition.[20]

Section 74 comes close to the definition proposed by the review by assuming that 'agreement by choice' cannot exist in the absence of freedom and capacity to make that choice.[21] To protect sexual autonomy in cases involving deception, the focus should be on whether the agreement was made as a result of manipulating C's decision to agree to sex. A person is unable to make an informed choice when her options are limited by deception. Thus, 'freedom' should inform 'choice'.

12 Richard Stanley Peters, 'Reason and Passion' in R F Dearden, P H Hirst and R S Peters (eds), *Education and the Development of Reason* (Routledge 2010) 161; D Pole, 'The Concept of Reason' in ibid 130; R P Wolff, *In Defence of Anarchism* (Harper & Row 1970) 14.

13 Home Office (n 3) para 2.10.5.

14 Ibid para 2.10.4.

15 Ibid para 2.10.5.

16 Ibid.

17 Law Commission, 'Consent to Sex' in Home Office, *Setting the Boundaries: Reforming the Law on Sex Offences* vol 2 (Home Office 2000) para 2.10.

18 Ibid para 5.1.

19 Home Office (n 3) para 2.10.5

20 Richard Card, Alisdair A Gillespie and Michael Hirst, *Sexual Offences* (Jordans 2008) 49.

21 Ibid.

Criminalising deception, mistake and non-disclosure

Liberal retributive principles require that only conduct which is blameworthy can legitimately be subject to state punishment.[22] Blameworthiness requires D to possess capacity for responsible agency. In other words, D knew what he was doing when he committed the offence, and exercised choice and a sufficient degree of control in doing so.[23] Blameworthiness, in the context of deception, is analysed in terms of the harm caused to C and the wrongfulness of obtaining sex by deception, non-disclosure or mistake. Harm deals with the degree to which the deceptive conduct causes, or risks causing a 'significant setback to another's interests'.[24] Determining wrongfulness involves examining the extent to which the criminal act involves the violation of a moral norm.[25] In the context of sex by deception, it is possible to violate C's sexual autonomy (wrongfulness) without causing her any harm.[26] It is argued that the wrongfulness of deception is that it may lead to a violation of sexual autonomy. The argument in favour of expanding the types of deception which negate consent is that rape should be understood as involving a violation of C's sexual autonomy.[27]

Sexual autonomy and the law of rape

Prior to the Act, the common law recognised only two types of deceptions which were capable of negating C's apparent consent: mistake as to the nature of the act[28] and mistake as to D's identity.[29] Mistakes falling outside these two fixed categories did not negate consent. The narrow approach of the common law is demonstrated in the case of *R v Linekar*.[30] C, a prostitute, agreed to sex with D on the understating that she would be paid £25. He in fact never paid and never intended to pay. D's conviction for rape was quashed. In *Linekar*, and other similar cases,[31] the evidence suggests that had C known the truth she would not have agreed to the sexual act, and in each case D had knowledge of this. Although these cases were correctly decided, in terms of the application of the law, it will be argued that in such cases, a duty of care ought to exist in sexual relationships and individuals owe a duty of responsibility when it comes to information regarding a material fact.

22 Stuart P Green, 'Lies, Rape, and Statutory Rape' in Austin Sarat (ed), *Law and Lies: Deception and Truth-Telling in the American Legal System* (Cambridge University Press 2015) 205.

23 Nicola Lacey and Hanna Packard, 'To Blame or to Forgive? Reconciling Punishment and Forgiveness in Criminal Justice' (2015) 35 Oxford Journal of Legal Studies 665.

24 Joel Feinberg, *The Moral Limits of the Criminal Law: Harm to Others* (Oxford University Press 1984) 31–36.

25 Green (n 22) 205.

26 John Gardner and Stephen Shute, 'The Wrongness of Rape' in J Gardner, *Offences and Defences: Selected Essay in the Philosophy of Criminal Law* (Oxford University Press 2007) 16.

27 Joan McGregor, 'Force, Consent, and the Reasonable Woman' in J Coleman and A Buchanan (eds), *Harm's Way: Essays in Honour of Joel Feinberg* (Cambridge University Press 1994) 231.

28 *R v Flattery* (1876–77) LR 2 QBD 410; *R v Williams* [1923] 1 KB 340.

29 *R v Elbekkay* [1995] Crim LR 163.

30 *R v Linekar* [1995] QB 250.

31 *Papadimitropoulos v The Queen* (1957) 98 CLR 249; *Bolduc and Bird v The Queen* (1967) 63 DLR (2d) 82.

Sexual autonomy

Both English law[32] and international law[33] recognise that rape is a violation of sexual autonomy. The next step is to determine whether deception relating to a material fact should constitute a violation of sexual autonomy which results in negating consent. The deception must have been 'material'. As Jonathan Herring rightly argues, 'she should not fall outside the law's protection simply because others do not agree with the reasons behind her sexual decisions'.[34] Rape law recognises the concept of autonomy in sexual relationships. However, the courts have yet to provide a definition of sexual autonomy. Jennifer Temkin argues that the overriding objective of the law of rape should be the protection of 'sexual choice', the individual's right to choose when and with whom to engage in sexual activity.[35] Anderson makes a similar argument by contending that sexual autonomy involves 'freedom from undesired sexual activity and freedom to engage in desired sexual activity'.[36]

To argue that deception relating to a material fact violates sexual autonomy and negates consent, it is important to highlight the value of sexual autonomy and the problems caused by its erosion. Sexual autonomy is limited by the rights of others and, as a result, it cannot entail the freedom to have sex whenever and with whomever one chooses.[37] It has been described as vital to the goals of securing women's equality with men and promoting a more general form of autonomy for women.[38] Thus, a special type of autonomy should be afforded to individuals in relation to their sexual activity.[39] It is argued that autonomy should refer to self-direction rather than to self-sufficiency.[40] The current law of rape does not allow individuals to achieve self-direction where their decision to agree to sex has been manipulated by deception. Deceptive conduct, relating to a material fact, influences an individual's agreement to sex by limiting the options available to her.

Autonomy has been described as comprising of two aspects: positive and negative.[41] Positive autonomy relates to an individual's right to decide when and with whom to engage in sex, whilst negative autonomy involves the right to refuse relations with others and have effect given to that refusal.[42] This description is misleading because of the tension between them. It is impossible to reconcile both positive and negative autonomy. While positive autonomy deals with an individual's right to choose, and the latter involves the right to refuse, the right to refuse should be the dominant aspect. However, this does

32 *R v Konzani* [2005] EWCA CRIM 706, [42].

33 *Prosecutor v Kunarac*, Case No IT-96–23-T & IT-96–23/1-T (International Criminal Tribunal for the Former Yugoslavia 22 February 2002) [440].

34 Jonathan Herring, *Criminal Law* (3rd edn, Palgrave 2015) 110.

35 Jennifer Temkin, 'Towards a Modern Law of Rape' (1982) 45 Modern Law Review 399, 401; see also Nicola Lacey where she provides a similar definition of sexual autonomy in 'Unspeakable Subjects, Impossible Rights: Sexuality, Integrity and Criminal Law' (1998) 11 Canadian Journal of Law Jurisprudence 47, 52.

36 Scott Anderson, 'On Sexual Obligation and Sexual Autonomy' (2013) 28 Hypatia 122, 133.

37 Vanessa E Munro, 'Sexual Autonomy' in M D Dubber and T Hornle (eds), *The Oxford Handbook of Criminal Law* (Oxford University Press 2014) 745.

38 Scott A Anderson, 'Prostitution and Sexual Autonomy: Making Sense of the Prohibition of Prostitution' (2002) 112 Ethics 748, 750–51.

39 Ibid 770.

40 Ibid.

41 Catherine Elliott and Claire de Than, 'The Case for a Rational Reconstruction of Consent in Criminal Law' (2007) 70(2) Modern Law Review 22, 231.

42 See Isiah Berlin, 'Two Concepts of Liberty' (originally published in 1958) in I Berlin, *Four Essays on Liberty* (Oxford University Press 1969).

not assist in resolving the issue of why inducing by lying about the fact that they were to have sex[43] or about D's identity negate consent whereas other deceptions, such as marital status, using birth control, carrying a sexually transmitted disease, or making false promises, such as D's commitment to entering into a long-term relationship, are not considered sufficiently serious to violate autonomy and negate consent. What it does show is that sexual autonomy is composed of a 'complex, multifarious collection of rights'[44] and that C's choice to refuse sex, based on a material fact, might be removed where her agreement is procured by deception, mistake or non-disclosure.

Choice and consent

Commentators such as Herring,[45] Vanessa Munro,[46] and Stephen Schulhofer[47] highlight the importance of treating sexual autonomy as a central principle. Lying and deceiving undermine sexual autonomy in a similar manner to force or the threat of force by limiting the options available to an individual.

The concept of choice, which forms part of the definition of consent in s 74, relates to notions of autonomy, liberty and responsibility.[48] In *R (on the application of F) v Director of Public Prosecutions*, Lord Judge CJ held that, 'choice' is crucial to the issue of 'consent'.[49] Unfortunately, his Lordship stated that the evidence relating to choice and freedom to make a choice should be approached in a 'broad common sense way'.[50] This dictum is similar to the direction in *Olugboja*,[51] which held that juries should '[apply] their combined good sense, experience and knowledge of human behaviour to all the relevant facts of that case', when distinguishing between consent and submission.[52] The *Olugboja* direction was criticised by the Home Office review for giving rise to confusion and uncertainty,[53] despite the fact that it removed the fixed category of circumstances which negated consent, and allowed individuals to set their own standards.[54] This can be contrasted with the current approach which fails to recognise an individual's own perceptions of her interests.

If choice presupposes that an individual has alternatives, her decision should be made on the basis of adequate information about each in order for that choice to be informed.[55] Choice presupposes two linked capabilities.[56] First, it involves a sense of self-awareness, whereby individuals are able to decide 'which of the accessible options will best realise [their] ideal, and thus suit [them]'.[57] This suggests that choice is a subjective

43 For example, fraudulent medical procedures.

44 Green (n 22) 205.

45 Jonathan Herring, 'Rape and the Definition of Consent' (2014) 26 National Law School of Indiana Review 6.

46 Vanessa E Munro, 'Constructing Consent: Legislating Freedom and Legitimating Constraints in the Expression of Sexual Autonomy' (2008) 41 Akron Law Review 923.

47 Stephen J Schulhofer, *Unwanted Sex: Culture of Intimidation and Failure of Law* (Harvard University Press 2000).

48 Gerald Dworkin, *The Theory and Practice of Autonomy* (Cambridge University Press 1988) 61.

49 *R (on the application of F) v Director of Public Prosecutions* [2013] EWHC 945 (Admin); [2014] QB 581; [2014] 2 WLR 190, [26].

50 Ibid [26].

51 *R v Olugboja* [1982] QB 320.

52 Ibid 332.

53 Home Office (n 3) para 2.2.3.

54 Simon Gardner, 'Appreciating *Olugboja*' (1996) 16(3) Legal Studies 275, 287.

55 David Ormerod, *Smith and Hogan's Criminal Law* (14th edn, Oxford University Press 2015) 823.

56 David Levine, *Wealth and Freedom: An Introduction to Political Economy* (Cambridge University Press 1995) 28.

57 Ibid.

state of mind. Applying this subjective facet to sexual autonomy implies that individuals are able to choose between the alternatives available and to distinguish between them based on preferences. For some it may be their preference to only have sexual relationships with those from a particular religious or socio-economic group. The second element of choice is autonomy, 'the sense of having an inner core capable of identifying wants and pursuing their satisfaction'.[58] The relationship between self-awareness and autonomy is best articulated by Robert Lindley who suggests that autonomy consists of two elements. Firstly, it involves an individual acting on reasons based on her own goals and purposes.[59] The second element requires freedom from external constraints such as being manipulated by others to do their will.[60] To protect sexual autonomy, the law of rape should safeguard an individual's right not to be manipulated by lies or deception. Deceiving another into agreeing to sex is an affront to autonomy.[61]

The impact of deception on sexual autonomy

Although moral philosophers tend to distinguish between lying and deceiving,[62] the courts distinguish between active deception,[63] non-disclosure[64] and mistake.[65] Here, it will be argued that a more holistic concept of deception is preferable, because it addresses whether the deception employed impacted on another's decision to agree to sex. The selectivity associated with the current approach has resulted in inconsistency and fails in its primary purpose of protecting sexual autonomy.[66]

Deception encompasses an unlimited variety of methods by which the deceiver creates false impressions in another's mind.[67] The term is used in this article to refer to the communication[68] of a message[69] that is intended to mislead.[70] Deception, unlike lying, need not be directed at a specific individual. Examples of indirect deceptions include where D removes his wedding ring to mislead others about his marital status or where he gives the impression that he is a footballer or film producer.[71] Deception includes actions, omissions, words and strategic silences.[72] It is recognised that deception is an element of many forms of acceptable social behaviours, such as tact, politeness or evasion.[73] In the context of sexual relationships, parties rarely disclose every potentially

58 Ibid.

59 Robert Lindley, *Autonomy (Issues in Political Theory)* (Palgrave Macmillan 1986) 6.

60 Ibid.

61 Larry Alexander and Emily Sherwin, 'Deception in Morality and Law' (2003) 22 Law and Philosophy 393, 432.

62 Immanuel Kant, 'Ethical Duties towards Others: Truthfulness' in *Lectures on Ethics,* L Infield (trans) (Harper Row 1963) 266; Sissela Bok, *Lying: Moral Choice in Public and Private Life* (Random House 1978); Albert Vrij, *Detecting Lies and Deceit: Pitfalls and Opportunities* (Wiley 2000); D Fallis, 'What is Lying?' (2009) 106 Journal of Philosophy 29.

63 *McNally (Justine)* (n 5).

64 *R v B* [2006] EWCA Crim 2945.

65 *Linekar* (n 30).

66 Home Office (n 3) para 2.10.4.

67 Alexander and Sherwin (n 61) 400.

68 'Communication' can be direct or indirect.

69 Words and conduct.

70 Stuart P Green, *Lying, Cheating, and Stealing: A Moral Theory of White-Collar Crime* (Oxford University Press 2007) 76.

71 *R v Melitti* [2001] EWCA Crim 1563.

72 Alexander and Sherwin (n 61) 400.

73 Jonathan E Adler, 'Lying Deceiving, or Falsely Implicating' (1997) 94 Journal of Philosophy 435, 435.

relevant detail.[74] There are a variety of methods by which one can tacitly mislead. Examples include clothing, cosmetics, religious symbols and removing one's wedding ring. Sexual autonomy is violated where an individual resorts to such deceptions to procure sex, knowing that his deception relates to a material fact. According to Adler, 'deception need not be intentional or voluntary, as lying must'.[75] However, he concedes that both lying and deception aim for the victim to believe falsely.[76] Thus, sexual autonomy should be deemed to have been violated where a lie or deception might cause another to be misled and results in her agreement to sex, provided this deception related to a material fact.

Where D intentionally misleads C about a material fact and manipulates her agreement to sex, despite being aware of C's mistake, C's consent should be vitiated by the deception. Applying the concept of relational autonomy, a duty must be imposed on individuals to disclose a fact which they know is material to others or correct mistakes relating to a material fact. The relevance of relational autonomy is that it recognises the importance of relational obligations and responsibilities[77] which are vital in the context of sexual activity.

The impact of lying on sexual autonomy

Although a universally accepted definition of lying does not exist, it has been defined as a 'statement made by one who does not believe it with the intention that someone else shall be led to believe it'.[78] This definition requires further clarification to fully distinguish it from deception. Lying involves a much narrower range of behaviours than deception generally.[79] Deception, as shown above, includes a variety of methods by which the deceiver produces false impressions in another's mind.[80] Unlike deceptions, lies require an assertion that 'we present ourselves as believing something while and through invoking (although not necessarily gaining) the trust of the one' to whom we assert.[81]

Lying is viewed as wrong for various reasons. Thomas Aquinas maintained that lying is contrary to the law of nature.[82] Immanuel Kant also viewed lying, which he defined as 'false assertion', to be 'directly opposed to the natural purposiveness of the speaker's capacity to communicate his thoughts', and the liar 'throws away and, as it were, annihilates his dignity as a human being'.[83] Kant illustrates the absolute character of the moral imperative not to lie by giving the example of lying to the murderer at the door who asks about the whereabouts of his intended victim.[84] Lying is wrong because it

74 Jed Rubenfeld, 'The Riddle of Rape-by-Deception and the Myth of Sexual Autonomy' (2013) 122 Yale Law Journal 1372, 1372.

75 Adler (n 73) 435.

76 Ibid.

77 Jonathan Herring, 'Relational Autonomy and Consent' in Alan Reed and Michael Bohlander (eds), *Consent: Domestic and Comparative Perspectives* (Routledge 2017) 27.

78 Arnold Isenberg, 'Deontology and the Ethics of Lying' in William Callaghan, Leigh Cauman, Sidney Mothersill et al (eds), *Aesthetics and Theory of Criticism: Selected Essays of Arnold Isenberg* (University of Chicago Press 1973) 248; for an alternative definition see Alexander and Sherwin (n 61).

79 Green (n 70) 77.

80 Alexander and Sherwin (n 61) 400.

81 D Simpson 'Lying, Liars and Language' (1992) 52 Philosophy and Phenomenological Research 623, 625.

82 Thomas Aquinas, *Summa Theologiae* vol 4 (McGraw-Hill 1972).

83 Immanuel Kant, *The Metaphysics of Morals,* Mary Gregor (trans) (Cambridge University Press 1996) 182.

84 Immanuel Kant, 'On Supposed Right to Tell Lies from Benevolent Motives' in T K Abbott (ed and trans), *Kant's Critique of Practical Reason and Other Works on the Theory of Ethics* (Dover Publications 1989) 36.

violates autonomy by forcing an individual to pursue the speaker's objectives, rather than her own preferences.[85]

Lying asserts what D believes to be false.[86] Deceptions, on the other hand, are invitations by D to accept as true his deceptive behaviour.[87] In his treatment of the false promise case, under the Formula of Humanity, Kant explained that the victim of the lie would not agree to being used to the advantage of the false promisor and 'cannot contain the end of this action in himself'.[88] Kant's treatment of the false promisor can be equally applied to all forms of deception, including Herring's example of the rogue who falsely proclaims his love,[89] to show that deceptions violate sexual autonomy. Consent should be deemed to be absent where an individual is deceived in relation to a material fact and agrees to sex because of that deception. Sexual autonomy is violated because there is lack of reciprocity between the parties. But for D's deception, C would not have agreed to sex. Applying an autonomy-based argument, which focuses on protecting the individual's ability to choose from a set of options, suggests that a successful lie or deception distorts the reasoning process of the individual deceived. It displaces her will and manipulates her action for the speaker's end.[90]

Deception and the current law of sexual offences

To illustrate the deficiency of the law of non-consensual sex in cases involving lying and deceiving, it is illuminating to consider the decision of the Crown Prosecution Service (CPS) not to prosecute undercover police officers who engaged in sexual relations with women. In *AJA & Others v Commissioner of Police for the Metropolis*,[91] the CPS halted criminal proceedings against six defendants who had been due to stand trial at Nottingham Crown Court on charges related to a conspiracy to sabotage a coal-fired power station at Ratcliffe-on-Soar. The CPS was concerned that Nottinghamshire Police had failed to comply with its pre-trial disclosure obligations relating mainly to the work of undercover police officer Mark Kennedy.[92] It later emerged that PC Kennedy had had at least one long-term, intimate sexual relationship with a woman involved with one of the groups he had infiltrated. Following this discovery, allegations concerning undercover police officers acting beyond their authorisation, or taking action which was authorised but should not have been, were reported in the media. It was claimed that several officers had intimate relationships with members of the groups they had infiltrated. One officer was said to have fathered a child in such a relationship before disappearing.[93]

In February 2013, the Home Affairs Committee invited the women involved to give evidence. One witness explained that:

85 D A Strauss, 'Persuasion, Autonomy, and Freedom of Expression' (1991) 91 Columbia Law Review 334, 355.

86 F A Siegler, 'Lying' (1966) 3 American Philosophical Quarterly 128, 130.

87 Green (n 70) 77.

88 Immanuel Kant, *The Moral Law,* H J Paton (trans) (Hutchinson & Co 1969) 91.

89 Herring (n 2) 511.

90 Kant (n 62); Alexander and Sherwin (n 61) 397; P J Griffiths, *Lying: An Augustinian Theology of Duplicity* (Brazos Press 2004); J E Mahon, 'A Definition of Deceiving' (2007) 21 International Journal of Applied Philosophy 181.

91 *AJA & Others v Commissioner of Police for the Metropolis* [2013] EWCA Civ 1342; [2014] 1 WLR 285.

92 *Ratcliffe-on-Soar Power Station (Operation Aeroscope) Disclosure* (Reference: 2011/000464, Final Report of an Independent Investigation by the Independent Police Complaints Commission 2012).

93 P Lewis, R Evans and S Pollack, 'Trauma of Spy's Girlfriend: "Like being Raped by the State"' *The Guardian* (London, 24 June 2013) <www.theguardian.com/uk/2013/jun/24/undercover-police-spy-girlfriend-child>.

> How it feels to me is that it is not having found out that your partner was lying about who they are; it is finding out that your most personal relationship was being controlled by the state without your knowledge.[94]

Another woman who had a child with an undercover police officer stated:

> We are psychologically damaged; it is like being raped by the state. We feel that we were sexually abused because none of us gave consent.[95]

It is, therefore, clear that the complainants would not have engaged in sex with the undercover police officers had they known their true identities. The conduct of the undercover police officers, with which the Home Affairs Committee[96] concerns itself, appear to both pre-date[97] and post-date the Act.[98] The CPS examined various leading authorities on the meaning of consent, as set out in the Act, and concluded that, in accordance with the principles established by case law, there was insufficient evidence to prosecute for rape because 'any deceptions in the circumstances of this case were not such as to vitiate consent'.[99]

The first case to be considered by the CPS was that of *Assange v Sweden*.[100] The Divisional Court was asked to determine, inter alia, whether the principle of dual criminality was met on the facts of the case. It was alleged that Assange had intercourse with C after deliberately creating a tear in the condom, or that he had removed it, knowing that C had only consented to sex on the condition that he wore a condom. According to the court, C was not deceived as to the 'nature or purpose of the act' as set out in s 76(2)(a) of the Act. In emphasising the limited scope of s 76, the court stated:

> [Section 76] deals simply with a conclusive presumption in the very limited circumstances to which it applies. If the conduct of the defendant is not within s 76, that may not preclude reliance on s 74. *R. v B* goes no further than deciding that failure to disclose HIV infection is not of itself relevant to consent under s 4.[101]

The issue of non-disclosure, which is a form of deception, was discussed in *B*:

> [The] fact that the defendant may not have disclosed his HIV status is not a matter which could in any way be relevant to the issue of consent under section 74 in relation to the sexual activity in this case.[102]

Thus, active non-disclosure of HIV does not vitiate consent. It is difficult to justify this position given the fact that D's HIV-positive status was a material fact in terms of C's decision to agree to sex.

94 House of Commons – Home Affairs Committee, *Undercover Policing: Interim Report* (HC 837, Thirteenth Report of Session 2012–13, TSO 2013) <www.publications.parliament.uk/pa/cm201213/cmselect/cmhaff/837/83702.htm>.

95 Ibid.

96 Ibid.

97 Mark Kennedy met Anna in 2005, when she was 21 years old.

98 Robert Lambert, using the undercover persona of Mark 'Bob' Robinson, formed a sexual relationship with a woman, who was unaware of his true identity at the time.

99 Crown Prosecution Service, 'Charging Decision Concerning MPS Special Demonstration Squad' (*Blog of the Crown Prosecution Service*, 21 August 2014) <http://blog.cps.gov.uk/2014/08/charging-decision-concerning-mps-special-demonstration-squad.html>.

100 *Assange* (n 5).

101 Ibid [90].

102 *R v B* (n 8) [21].

The second case considered was *R (on the application of F)*.[103] The facts of this case were similar to those of *Assange*, in that C, who did not wish to become pregnant, consented to sex with her husband on the condition that he would withdraw his penis before ejaculating within her vagina. During intercourse, shortly after penetration, D informed her that he would be ejaculating within her vagina 'because you are my wife and I'll do it if I want'.[104] He ejaculated before she could say or do anything about it. The CPS decided not to prosecute D for rape because the evidence would be insufficient to establish a realistic prospect of conviction. C sought judicial review of that decision. Lord Judge CJ stated that, at the time of reaching the decision not to prosecute, the Crown Prosecutor had not had the benefit of the decision of the Divisional Court in *Assange*:

> What the *Assange* case underlines is that 'choice' is crucial to the issue of 'consent' … The evidence relating to 'choice' and the 'freedom' to make any particular choice must be approached in a broad common sense way. If before penetration began the interested party had made up his mind that he would penetrate and ejaculate within the claimant's vagina, or even, because 'penetration is a continuing act from entry to withdrawal' (see section 79(2) of the 2003 Act) he decided that he would not withdraw at all, just because he deemed the claimant subservient to his control, she was deprived of choice relating to the crucial feature on which her original consent to sexual intercourse was based. Accordingly, her consent was negated. Contrary to her wishes, and knowing that she would not have consented, and did not consent to penetration or the continuation of penetration if she had any inkling of his intention, he deliberately ejaculated within her vagina. In law, this combination of circumstances falls within the statutory definition of rape.[105]

It is important to note that in *Assange* and *R (on the application of F)*, the agreement to sex was based on a stipulated condition. Rogers criticises *Assange* because liability for rape and other non-consensual sexual offences should only arise where C has not been willing 'to be used for the sexual gratification of another in a way that shows regard to one's own sexual preferences'.[106] This analysis endorses the fact that a commitment to sexual autonomy should take into account the significance of the sexual act for C, and her state of mind at the time of agreeing to engage in sex.

The final case considered was *R v McNally*,[107] in which D, a female who posed as a boy called 'Scott' on an online social media site, deceived C as to her gender. D pleaded guilty to six counts of assault by penetration, contrary to s 2 of the Act. D appealed on the grounds that she should not have been convicted at trial because the elements of the offence had not been made out. She argued that deception as to gender could not vitiate consent, being deception as to a quality or attribute, as stated in *B*.[108] D contended that she had not been advised as to the relevant legal elements when deciding to plead guilty. D also appealed against sentence. The Court of Appeal dismissed her appeal against conviction but allowed her appeal against sentence.

In relation to the first ground for appeal, the court held that 'some deceptions (such as, for example, in relation to wealth) will obviously not be sufficient to vitiate

103 *R (on the application of F) v DPP* [2013] EWHC 945 (Admin).

104 Ibid.

105 *R (on the application of F)* (n 103) [26].

106 Jonathan Rogers, 'The Effect of "Deception" in the Sexual Offences Act 2003' (2013) 4 Archbold Review 7.

107 *R v McNally (Justine)* (n 5)

108 *R (on the application of F)* (n 103).

consent'.[109] Thus, it is settled law that not all deceptions will negate consent. It is unclear whether this dictum provides any indication as to the scope of s 74.[110]

Two issues stem from the court's decision to limit the types of deceptions which negate consent. Firstly, the court has failed to recognise the importance of a material fact to the individual complainant. For some it is important that their sexual partner is of a certain religion or free from HIV, while the same facts might not be relevant to others. The concern with framing D's liability in terms of 'active deception' is that consent under s 74 would only be vitiated where D lied to C about a fact recognised by the courts. Consent will not be negated where D withholds a material fact or where he takes advantage of a mistake, despite the fact that C would not have agreed had she known the truth. Secondly, the court did not provide an explanation as to why only gender deception could negate consent, while other categories such as wealth would not vitiate consent. The danger of this approach is that it considers certain deceptions as too trivial, irrespective of their importance to the individual and risks giving rise to discrimination.[111]

Returning to the case of the undercover police officers, the officers withheld a material fact relevant to the complainants, namely that they were undercover police officers. Although, under the current law, non-disclosure, a form of deception, may not negate consent, the undercover officers, to perpetuate their tacit deception, would undoubtedly have had to engage in active deception or lying. Thus, in some cases which appear to involve only tacit deception, D might have also resorted to lying to maintain the deception. This analysis does not suggest that being a police officer is akin to gender. The argument put forward in this article is that deception relating to any material fact, whether it involves religion, gender or wealth, could negate consent.

The current distinction between lying and deceiving is applied without justification and results in the failure of the law to provide an effective framework for the protection of sexual autonomy. While it is accepted that in the context of the criminal law in practice, individuals may be prevented from setting their own boundaries of acceptability, an autonomous individual should be 'exempt from arbitrary control, un-coerced and unrestricted'.[112] Individuals should be directed by considerations, desires, conditions and characteristics that are not simply imposed externally upon them.[113]

Proposal for reform: removing the distinction between lying and deceiving

A tradition in ethics maintains that lying is a significantly worse form of behaviour than deception.[114] A sharp distinction between lying and deception is drawn by deontological theorists such as Kant, who focused on the nature of false assertions. Hence, there is no

109 *R v McNally (Justine)* (n 5) [25].

110 Karl Laird, 'Rapist or Rogue? Deception, Consent and the Sexual Offences Act 2003' [2014] Criminal Law Review 492, 506.

111 Gender fraud cases include *R v Barker (Gemma Louise)* [2012] EWCA Crim 1593; *R v Chris Wilson* [2013] unreported; *R v Kyran Lee (Mason)* [2015] unreported; *R v Gayle Newland* [2017] unreported (the defendant's actual name is Gail, but due to a typographical error in a court report she has always been referred to as Gayle); for a detailed critique of the court's decision in *R v McNally (Justine)* (n 5), see Alex Sharpe, 'Criminalising Sexual Intimacy: Transgender Defendants and the Legal Construction of Non-consent' [2014] Criminal Law Review 207.

112 Matti Wiberg, 'Political Autonomy: Ambiguities and Clarifications' in M Suksi (ed), *Autonomy: Applications and Implications* (Kluwer Law International 1998) 44.

113 John Christman, 'Constructing the Inner Citadel: Recent Work on the Concept of Autonomy' (1988) 99(1) Ethics 109, 114.

114 Adler (n 73) 435.

wrong in the act of packing luggage in the presence of others to create the impression that one is about to leave town.[115] While those who observe this conduct may form a false belief, they 'have no right to expect that my action will express my real mind'.[116] Although the deceiver has succeeded in deceiving others, Kant insisted that: 'I have not lied to them, for I have not expressed an opinion.'[117] It is submitted that Kant's analysis of his example should not apply to sexual offences as the act of packing was intended to create a false belief that he was embarking on a journey. By making a false statement, the liar is indicating that he is expressing an opinion.[118] This form of deception, which involves taking advantage of another's mistake, is capable of violating sexual autonomy if an individual is mistaken about a material fact and the deceiver, rather than correcting the mistake, takes advantage of this and uses C for his own gratification.

Roderick Chisholm and Thomas Feehan ask, 'why is lying thought to be worse . . . than other types of intended deception?'[119] The authors' answer is that only lying involves assertion and 'lying, unlike the other types of intended deception, is essentially a breach of faith'.[120] Applying this reasoning to Kant's example of the packed bags, the authors argue that he has not invited his friends to assume that he is about to make a journey.[121] Chisholm and Feehan (as well as Kant) fail to take into account the impact of deception on the individual's decision, as well as the fact that the deceiver has deliberately packed the bags for the purpose of misleading his friends. Deceptions may not involve an assertion, but they can amount to a breach of faith where the deceiver acts on the other's mistake. A prostitute who agrees to sex in the mistaken belief that she will be paid has had her decision manipulated by this deception.[122] While it is true that she has agreed to sex and has not been misled as to the mechanics of the act, her agreement is based on the understanding that she would be paid for sex. Had she known the truth she would not have agreed to sex. Deception, therefore, uses an individual's own decision-making powers against them.[123] Alan Wertheimer notes that 'current social norms may understate the seriousness of sexual deception'.[124] He suggests that sexual consent may be vitiated by deception about one's marital status, an affair with a partner's sister, one's views on contentious moral issues like abortion, one's feelings, or one's intentions to marry.[125]

Failure to protect sexual autonomy in cases involving mistake on C's part appears to hold individuals responsible for making an inference resulting from another's actions. This appears analogous to the concept of contributory negligence. The individual's ability to exercise self-determination in pursuing her own interests implies that others have a duty to honour this attempt, as she must honour that of others. Taking advantage of another's mistake, or deliberately engineering or procuring that mistake, is a violation of sexual autonomy because the person's decision to engage in sex has been influenced by the conduct of the deceiver. Kant's distinction between lying and deceiving is presumably

115 Kant (n 62) 226.

116 Ibid.

117 Ibid.

118 Roderick M Chisholm and Thomas D Feehan, 'The Intent to Deceive' (1977) 74 Journal of Philosophy 143, 149.

119 Ibid 153.

120 Ibid.

121 Ibid.

122 *Linekar* (n 30).

123 Herring (n 2).

124 Alan Wertheimer, *Consent to Sexual Relations* (Cambridge University Press 2003) 199.

125 Ibid.

based on the idea that individuals are rational, autonomous beings and ultimately responsible for their inferences.[126] However, the distinction cannot be justified where the end result is obtaining sex. Whether sex is obtained by lying or deceiving, sexual autonomy has been violated.

Returning to the case of the undercover police officers, it is irrelevant whether the deception as to their true identity was active or tacit. The issue for the courts should be whether withholding information manipulated C's sexual decision. Given the complainants' evidence to the Home Affairs Committee, it is doubtful they would have agreed to sex with police officers.

Non-disclosure and mistake should nullify consent because it prevents individuals from setting their own standards with regards to the characteristics of their sexual partners. Consent should be deemed valid where C is mistaken about D's attributes, such as his marital status, religious affiliations and wealth. Where D withholds information relating to a material fact, and he does so for the purpose of manipulating her decision to have sex, C's consent should be considered to have been vitiated by his non-disclosure. The focus should be on the impact the deceptive conduct had on the individual's choice. In *R v Cuerrier*,[127] the Supreme Court of Canada held that '[t]he dishonest action or behaviour must be related to the obtaining of consent to engage in sexual intercourse'.[128] This should consist of either deliberate deceit regarding D's HIV status or non-disclosure of that status.[129] Thus, without disclosure of HIV status, consent cannot be freely given. According to the court, 'consent cannot be simply to have sexual intercourse . . . it must be to have intercourse with a partner who is HIV-positive'.[130] This reasoning can be applied to any material fact. The issue is whether D manipulated C's sexual choice by taking advantage of her mistake.

A more recent Canadian case, with facts similar to *Assange*, is *R v Hutchinson*.[131] C agreed to sexual activity with D, insisting that he use a condom to prevent conception. D deliberately poked holes in the condoms to get C pregnant without her permission. He succeeded in getting C pregnant, and this resulted in her having an abortion, which gave rise to complications. The trial judge found that the complainant had not consented to unprotected sex and convicted H of sexual assault. On appeal, the majority upheld the conviction on the basis that condom protection was an 'essential feature' of the sexual activity, and therefore the complainant did not consent to the 'sexual activity in question'. In relation to the concept of 'harm' caused by D's deception, the court concluded that:

> '[H]arm' does not encompass only bodily harm in the traditional sense of that term; it includes at least the sorts of profound changes in a woman's body – changes that may be welcomed or changes that a woman may choose not to accept – resulting from pregnancy. Depriving a woman of the choice whether to become pregnant or increasing the risk of pregnancy is equally serious as a 'significant risk of serious bodily harm' within the meaning of *Cuerrier*, and therefore suffices to establish fraud vitiating consent.[132]

126 Adler (n 73) 444.

127 *R v Cuerrier* [1998] 2 SCR 371.

128 Ibid.

129 Ibid.

130 Ibid.

131 *R v Hutchinson* (2013) 105 WCB (2d) 806 (a decision of the Nova Scotia Court of Appeal).

132 Ibid [70].

Hutchinson both actively deceived C and took advantage of her mistake relating to a material fact, namely, that he would use a condom to avoid an unwanted pregnancy. The court held that there was no consent because protected sex was an essential feature of the proposed sexual activity and not a separate component from what C was consenting to.[133] A fact which is material to C should, therefore, be an essential feature of the sexual activity.

The law's deficiency and inconsistent approach to responding adequately in cases involving non-disclosure of a material fact can be illustrated by using two examples[134] in which deception was used. In *R v Richardson*,[135] it was held that a dentist had not assaulted her patients merely because she treated them whilst suspended from practice. Allowing the appeal, it was held that an assault could only occur where consent was given in the mistaken belief that the dentist was other than she truly was. Although the complainants were unaware that D was no longer qualified to practise, they were fully aware of her identity. D's non-disclosure, about a material fact such as medical qualifications, would not go to the nature or purpose of the act. A different outcome was reached in *R v Tabassum*,[136] where several complainants allowed D to examine their breasts on the basis of a false representation that he was medically qualified and conducting a survey into breast cancer. The evidence showed that, in fact, he was attempting to develop software to increase his chances of being accepted into a medical school. In dismissing his appeal, the court concentrated on the fact that D had deceived the complainants as to his medical qualifications.[137]

While both the above cases concern D's lack of medical qualifications, a distinction can be drawn: the complainants in *Richardson* were mistaken due to D's deliberate non-disclosure, whereas Tabassum actively deceived the complainants. In both cases, the defendants carried out deception in order to use the complainants for their own gain; Richardson so that she could continue to practise as a dentist, and Tabassum so that he could develop his software. The argument made in this paper is that this distinction is arbitrary, ethically suspect, has no regard for personal autonomy and should not form the basis of sexual offence laws.

A potential objection to this proposal might be that it requires individuals from a sexual minority to disclose intimate personal details regarding their sexual past or sexual orientation. For example, a bisexual man who does not disclose his sexual orientation could be criminalised under the proposed recommendation. Similarly, a post-operative transsexual who fails to disclose this fact could also be criminally liable. It could be argued that on policy grounds non-disclosure of these facts should constitute an exception to the proposal advanced in this article. However, if the primary aim is the protection of sexual autonomy, as this article has argued, distinctions should not be made between different material facts.

Practical steps to protecting sexual autonomy

It has been argued that in protecting sexual autonomy where C's agreement is obtained by deception, active or tacit, criminal law should allow for the potential that any material fact may negate consent by removing the distinction between lying and deceiving. The

133 Ibid [98].

134 It is acknowledged that neither case deals with the Sexual Offences Act 2003.

135 *R v Richardson* [1999] QB 444.

136 *R v Tabassum* [2000] 2 Cr App R 328.

137 Ibid.

deception, lie or mistake must relate to a material fact[138] and must have induced C to agree to sex. However, the deception need not be directed at C. Similarly, C's mistake in relation to a material fact about D need not be created by D. Since the sheer use of another is morally wrong, D must not only respect C's autonomy but must also take reasonable steps to allow C to exercise her autonomy and ensure that C is free from deceptions and manipulation to provide consent.[139] An example of a deception which is not specifically aimed at a complainant is where D, who is married, creates a profile on an online-dating website. He 'meets' with C online who explains that she is seeking a long-term relationship and that she does not wish to engage in relationships with married men. They decide to meet in person and, prior to their meeting, D removes his wedding ring. If C agrees to sex, her consent is vitiated by D's deception due to various factors. First, D's representation about his marital status is false and he knows that it is false. Secondly, the deception can take the form of conduct or omission. An omission occurs where D fails to correct C's mistake about a material fact. Thirdly, D's deception, or lie, must be one 'of fact'.[140] This requirement also covers promises made by D which invite reliance upon and which are made in order to induce C into agreeing to sex.[141] Fourthly, the deception, unlike a lie, need not be directed at C. This can be achieved either directly or indirectly. An example of indirect deception is where D conveys information to a third party with the intention that the third party will convey the information to C.[142] Fifthly, the deception, lie or mistake must have induced C's agreement. To cover situations where the deception was an inducement which was not actively present in C's mind at the time at which she agreed to sex,[143] inducement includes circumstances where D deliberately withheld a fact which may be material to C. This requirement imposes a duty on D to disclose information which is material if he reasonably believes that it will influence her agreement.[144] In the context of contract law, English law does not require the representation to take the form of words.[145] Representation can be made by conduct or can be inferred from the facts and circumstances of the case. In sexual relationships, information can be conveyed as much by conduct as by words. Thus, removal of a wedding ring or wearing religious symbols ought to be used as evidence that C was induced by D's deception. Similarly, active concealment or non-disclosure of a material fact ought to negate consent. If a relationship with an insurance company gives rise to *uberrimae fidei*,[146] it seems extraordinary that sexual relationships are prevented from being granted similar status.[147] Criminal law does recognise liability for non-disclosure in the context of fraud, which imposes a duty to 'disclose to another person information which he is under a legal duty to disclose'. [148] No such duty exists in relation to sexual activity between adults of sound mind.[149] Imposing a duty of disclosure and not taking

138 According to the individual complainant, is an essential feature of the sexual activity.

139 Herring (n 77) 32.

140 *People v Evans* (1975) 85 Misc 2d 1088.

141 *Linekar* (n 30).

142 *Commercial Banking of Sydney v RH Brown and Co* [1972] 2 Lloyd's Rep 360.

143 *Horsfall v Thomas* (1862) 1 H & C 90.

144 The *mens rea* requirement for non-consensual sexual offences ensures that D will not be prosecuted when he did not know that C would regard a particular fact fundamental to her consent.

145 *Spice Girls Ltd v Aprilia World Service BV* [2000] EWHC Ch 140.

146 Utmost good faith.

147 Peter Alldridge, 'Sex, Lies and the Criminal Law' (1993) 44(3) Northern Ireland Legal Quarterly 250, 261.

148 S 3 of the Fraud Act 2006.

149 S 34 of the Sexual Offences Act 2003 deals with procuring sexual activity by inducement, threat or deception with a person with a mental disorder.

advantage of C's autonomy in sexual relationships requires accepting that individuals are entitled to engage in a cooperative and mutually beneficial relationship.[150]

Conclusion

As shown by the court decisions above, the law does not allow individuals the freedom to act according to their own sexual preferences. The criminal law should protect the right to sexual autonomy and sexual integrity.[151] Lying and the broader concept of deceiving have the same potential to negative consent. The method of deception employed should thus be irrelevant, provided the deception or lie relates to a material fact and results in manipulating another's sexual choice. Both lying and deceiving negatively impact on sexual autonomy by limiting options with the consequence that an individual cannot be said to have consented to sex. The issue of materiality is significant, but is essentially a question of fact to be determined in each case as opposed to a generic question of law.

Although this proposal may seem radical, it could in fact be accommodated within the existing law of sexual offences as it does not entail rejecting the definition of consent in s 74 of the Act. Any suggestion to introduce a lesser, distinct offence designed to capture deceptive forms of sexual coercion should be resisted as this obscures a vital principle: a lesser offence ignores the fact that deception intended to limit sexual choice negates consent. The proposed approach is akin to the *Olugboja* direction in that the jury would be required to focus on the state of mind of the individual. Widening the law in the manner suggested surely recognises that all forms of deception have the potential to violate sexual autonomy. Whether any given fact or factor is material is case-specific; so too is the reason why the defendant behaved as he did. It should not be forgotten that the defendant would also have to satisfy the *mens rea* of the offence. Complete autonomy is impossible. However, the concept developed in this article involves recognising that individuals are autonomous beings worthy of respect. Implicit in this is the right not to be deceived. This concept is based on the premise that individuals have legitimate interests capable of being pursued, which others should respect and which the criminal law should protect.

150 Herring (n 77) 32.

151 The difference between the right to sexual autonomy and to sexual integrity is discussed in Lacey (n 35).

NILQ 70(2): 221–239

Form and substance in contract damages

DAVID MCLAUCHLAN*

Victoria University of Wellington

Abstract

This article discusses the role of form and substance in the modern law of contract both generally and with specific reference to the law of damages for breach of contract and, in particular, the decisions of the UK Supreme Court in Swynson Ltd v Lowick Rose LLP *[2017] UKSC 32 and* Fulton Shipping Inc of Panama v Globalia Business Travel SAU (The New Flamenco) *[2017] UKSC 43. Although it was probably true to say when Atiyah and Summers wrote in* Form and Substance in Anglo-American Law *over 30 years ago that 'the English law of contractual damages continues to be treated by judges and writers as governed by highly formal rules', it would be wrong to describe the reasoning employed by judges in modern times when explaining, refining and applying these rules as highly formal. Particularly in appellate decisions, judicial reasoning is usually an amalgam of what the authors would describe as formal and substantive considerations. Indeed, the formal reason for supporting a decision may be preferred precisely because it provides the just or most convenient solution to the dispute, as in* Swynson v Lowick Rose. *In that case the Supreme Court overturned the decision of the majority of the Court of Appeal that denial of the damages claimed 'would be a triumph of form over substance', preferring the view of the dissenting judge who said that 'the form here is the substance'. And, while the decision in* The New Flamenco *appears at first sight to rest on formal, arguably formalistic, reasoning, a closer reading reveals that substantive considerations influenced the outcome of the appeal.*

Keywords: contract; damages; mitigation; avoided loss; causation; form and substance.

1 Introduction

In the period 21–24 November 2016 the UK Supreme Court heard together two important cases involving the assessment of damages for breach of contract, namely, *Swynson Ltd v Lowick Rose LLP* (*Swynson*)[1] and *Fulton Shipping Inc of Panama v Globalia Business Travel SAU* (*The New Flamenco*).[2] Although the facts of the cases were very different, they did raise a broadly similar issue as to whether the 'loss' in respect of which damages were claimed had been avoided: in *Swynson* by the post-breach conduct of a third party and in *The New Flamenco* by the post-breach conduct of the claimants themselves. In both cases the

* Professor of Law, Victoria University of Wellington; Professorial Fellow, University of Melbourne; Honorary Professor, University of Queensland. Thanks to Andrew Summers for his helpful comments on an earlier version of this article. Of course, the usual caveat applies.

1 [2017] UKSC 32, [2018] AC 313.

2 [2017] UKSC 43, [2017] 1 WLR 2581.

decisions of the Court of Appeal, in which Longmore LJ had delivered the leading judgments, were unanimously reversed by the Supreme Court. In *Swynson* the Court of Appeal had upheld the trial judge's award in favour of the claimants, whereas in *The New Flamenco* it had overturned the trial judge's award. Thus, the outcome of the litigation was that in *Swynson* the conduct of the third party was found to have extinguished the loss, but in *The New Flamenco* it was found that the conduct of the claimants had not done so.

A curious feature of the cases is that, although they were heard together, the decision in *The New Flamenco* was handed down nearly three months after that in *Swynson*.[3] This seems surprising at first sight because the latter raised a wider, and on the surface more difficult, range of issues. Thus, the claimants sought to uphold the award on a number of alternative bases, including the principle of 'transferred loss', unjust enrichment and equitable subrogation, and three full judgments making much the same points and running to a total of 121 paragraphs were delivered. By contrast, in *The New Flamenco* there was essentially one issue only and the case was resolved by a single judgment delivered by Lord Clarke which, after reciting the facts and the lower courts' reasoning, comprised a mere seven paragraphs of original analysis. It might therefore have been expected that the decisions would be handed down at least contemporaneously. It is, of course, dangerous to speculate on the reasons why this did not occur. One simple explanation is that in *Swynson* their Lordships regarded the correct outcome as relatively straightforward and that three of them preferred to write their own judgments, whereas in *The New Flamenco* the issue was initially seen as more difficult and requiring further deliberation, although eventually a consensus was reached which it was thought did not need a lengthy explanation or analysis of the authorities. Another explanation is that the judges in *The New Flamenco* were in fact deeply divided, not necessarily as to whether to allow the appeal but rather as to the grounds for doing so, and that, for one reason or another, when it was felt that delivery of judgment could be delayed no longer, it was decided that the case could be disposed of satisfactorily in a single and 'economically' reasoned judgment which Lord Clarke agreed to prepare. Having watched the whole of the argument online, I am inclined to think that the latter explanation may not be too wide of the mark, but naturally we will never know.

The main purpose of this article is to discuss the two decisions against the background of the renewed interest recently in the relationship between form and substance in the law of obligations, an interest sparked in large part by the rise in the number of cases in which the common law courts have exhibited impatience with perceived 'formalistic' or 'legalistic' distinctions that are not in accordance with 'economic reality' or otherwise ignore the 'substance' of the matter. This topic was the theme of the important Obligations IX conference held in Melbourne in July 2018. Included among the issues discussed were: to what extent do the Commonwealth courts now place more emphasis in their reasoning on the substantive considerations that underlie the rules and doctrines of private law and hence lead them to exhibit a greater preparedness to change or qualify those rules and doctrines where that is thought to be justified; what precisely is the nature of the distinction between formal and substantive reasoning; and does it provide worthwhile insights into the manner in which modern, especially appellate, courts resolve private law disputes? My modest contribution to the discussions, of which this article is a revised version, was to suggest that, although the two cases I have highlighted contain elements of formal reasoning, they do tend to support my view that searching for distinctions between formal and substantive reasoning in the authorities that make up the

3 The judgment in *Swynson* was handed down on 11 April 2017. The judgment in *The New Flamenco* was handed down on 28 June 2017.

modern law of contract, while an interesting endeavour, provides limited insights into the judicial decision-making process nowadays. The conference programme succinctly described formal reasoning as 'the application of rules without reference to the justifications that underlie them' and substantive reasoning as reasoning that makes 'direct reference to considerations of purpose, justice or convenience', but it is almost a truism to say that very often the two types of reasoning merge into one another. Indeed, Atiyah and Summers in their monumental yet now rather dated work on the subject, which includes many complex refinements of what constitutes formal and substantive reasoning in the first place, concede that '[a] formal reason usually incorporates or reflects substantive reasoning. Thus it is an admixture of certain formal attributes on the one hand, and substantive reasoning on the other.'[4] Consequently, a formal reason that is inappropriate because, for example, it ignores substantive considerations is regarded as *formalistic* and hence a bad reason.[5]

Although it was probably true to say when Atiyah and Summers wrote over 30 years ago that 'the English law of contractual damages continues to be treated by judges and writers as governed by highly formal rules',[6] such as the rules relating to remoteness of damage and mitigation, it would be wrong to describe the reasoning employed by judges in modern times when explaining, refining and applying these rules as 'highly formal'. Particularly in appellate decisions, judicial reasoning is usually an amalgam of what the authors would describe as formal and substantive considerations.[7] Indeed, the formal reason for supporting a decision may be preferred precisely because it provides the just or most convenient solution to the dispute. As we shall see, the latter point is illustrated by the decision in *Swynson*. The Supreme Court overturned the decision of the majority of the Court of Appeal that denial of the damages claimed 'would be a triumph of form over substance'.[8] Their Lordships preferred the view of the dissenting judge who said that 'the form here *is* the substance'.[9] Furthermore, as we shall also see, although categorisation of the judgment in *The New Flamenco* is more problematic because the main reasons for allowing the appeal seem at first sight to be highly formal and, in one important respect, unsatisfactorily formalistic, a closer reading reveals that the outcome of the appeal was influenced by substantive considerations.

However, before discussing the cases, it may be helpful to provide some wider context by considering the role of form and substance in the modern law of contract as a whole.

2 Form and substance in the modern law of contract

In his 1986 essay 'Form and Substance in Contract Law', Professor Atiyah argued that:[10]

> . . . everyone must be aware of the fact that the power of formal reasons in contract law and, indeed, perhaps in all of the law, has been declining in recent years. More and more often the courts seem willing to unpick the transaction, to

4 P S Atiyah and Robert S Summers, *Form and Substance in Anglo-American Law* (Clarendon Press 1987) 2.

5 Ibid 29 ('formalistic reasoning involves a failure to take substantive considerations into account *when they ought to be taken into account*').

6 Ibid 84.

7 See, eg, *Ruxley Electronics and Construction Ltd v Forsyth* [1996] 1 AC 344 (HL); *Farley v Skinner* [2001] UKHL 49, [2002] 2 AC 732; *Golden Strait Corpn v Nippon Yusen Kubishika Kaisha (The Golden Victory)* [2007] UKHL 12, [2007] 2 AC 35; *Transfield Shipping Inc v Mercator Shipping Inc (The Achilleas)* [2008] UKHL 48, [2009] 1 AC 61; *Bunge SA v Nidera BV* [2015] UKSC 43, [2015] 2 Lloyd's Rep 469; *Cavendish Square Holding BV v Makdessi* [2015] UKSC 67, [2016] AC 1172; *One Step (Support) Ltd v Morris-Garner* [2018] UKSC 20, [2018] 2 WLR 1353.

8 [2015] EWCA Civ 629, [2016] 1 WLR 1045 [16] (Longmore LJ).

9 Ibid [36] (Davis LJ).

10 Essay 6 in P S Atiyah, *Essays on Contract* (Oxford University Press 1986) 116.

> open it up, as it were, and go behind the formal reasons, and look at the substantive reasons for the creation or negativing of obligations. In some instances, these developments in the law might seem to amount to a mere narrowing of the limits within which formal reasons are treated as conclusive. But in other cases it is hard to avoid the conclusion that the formal reasons themselves have been jettisoned, and the issues treated as though they were completely at large. The virtual disappearance in modern English law of the parol evidence rule seems one clear example of the emasculation of a formal reason in contract cases.

Some 10 years later, Lord Steyn, writing extra-judicially, expressed a similar view. In opening his argument that the fundamental purpose of the law of contract is to give effect to the reasonable expectations of contracting parties, he said that '[t]he modern view is that the reason for a rule is important' and that '[t]he rule ought to apply where reason requires it, and no further'.[11] Thus, it is an important and legitimate part of the judicial function, particularly in appellate courts, to identify the purpose of legal rules and re-examine whether, as usually formulated, they fulfil their purpose, taking into account that 'simple fairness ought to be the basis of every legal rule and, in a common law case, that the presumption in favour of the fair solution is powerful'.[12]

The above observations echo a remark, not cited by either author (or indeed Atiyah and Summers), by Lord Wilberforce in the 1980 case of *National Carriers Ltd v Panalpina (Northern) Ltd* (*Panalpina*). His Lordship said:[13]

> I think that the movement of the law of contract is away from a rigid theory of autonomy towards the discovery — or I do not hesitate to say imposition — by the courts of just solutions, which can be ascribed to reasonable men in the position of the parties.

Many examples of the movement his Lordship described can be found scattered throughout the case law during the nearly 40 years since he wrote.[14] We have witnessed the abandonment or modification of several of the formal doctrines of the law of contract established during the nineteenth century and the first half of the twentieth

11 Johan Steyn, 'Contract Law: Fulfilling the Reasonable Expectations of Honest Men' (1997) 113 Law Quarterly Review 433, 433.

12 Ibid.

13 [1981] AC 675, 696.

14 Of course, there is nothing entirely novel in the idea that the law of contract seeks to find just solutions that can be ascribed to reasonable people in the position of the parties, for it is trite learning that one of the major purposes of the law of contract is to protect the reasonable expectations of the parties. As Professor Corbin observed at the outset of his treatise (A L Corbin, *Corbin on Contracts* vol 1 (rev edn, West Publishing Co 1960) §1, p 2)): 'That portion of the field of law that is classified and described as the law of contracts attempts the realisation of reasonable expectations that have been induced by the making of a promise. Doubtless, this is not the only purpose by which men have been motivated in creating the law of contracts; but it is believed to be the main underlying purpose, and it is believed that an understanding of many of the existing rules and a determination of their effectiveness require a lively consciousness of this underlying purpose.' Nevertheless, this purpose has had an even more pervasive influence in the evolution of the modern law of contract. It is reflected, for example, in the cases that have imposed duties of good faith in the performance of contracts (see *Yam Seng Pte v International Trade Corp Ltd* [2013] EWHC 111 (QB), [2013] 1 Lloyd's Rep 526) and, strikingly, in others where the courts have reviewed the exercise of contractual discretions and decision-making powers essentially on administrative law grounds of review, in particular, the *Wednesbury* rationality test, so that the contractual decision-maker is required to exercise its power honestly and genuinely in conformity with the nature and purpose of the contract, and must exclude considerations that are irrelevant in the light of that purpose, the terms of the contract and the context. See *Braganza v BP Shipping Ltd* [2015] UKSC 17, [2015] 1 WLR 1661 and *Waaler v Hounslow London Borough Council* [2017] EWCA Civ 45, [2017] 1 WLR 2817.

century. Let us take the relatively small area that is the main subject of this article, the law of damages. Here we have seen the refinement, much discussed in the academic literature, of the law relating to remoteness of damage. Atiyah and Summers argued that '[e]ven with respect to problems which are inherently difficult to force into a framework of formal rules—such as the foreseeability principle of *Hadley v Baxendale*—English courts still try to formulate the law in terms of specific formal rules, and most English judges try (if, as some may think, rather futilely) to apply these rules formally',[15] citing in support the judgments of the House of Lords in *Koufos v C Czarnikow Ltd* (*The Heron II*).[16] However, that formal approach has since been rejected by a majority of their Lordships who, in *Transfield Shipping Inc v Mercator Shipping Inc* (*The Achilleas*),[17] endorsed an agreement-centred approach to remoteness under which a loss may be considered too remote, even if it was of the type or kind that would have been within the reasonable contemplation of the parties as a not unlikely result of the breach. Although, as Lord Hoffmann conceded, satisfying the latter requirement will usually suffice, it will not always do so. Defendants will escape liability for foreseeable loss if it cannot reasonably be inferred that they accepted responsibility for that loss.[18]

Other developments have included the rejection of the rule that damages cannot be recovered for late payment of a debt,[19] substantial modification of the rule that damages are not available for mental distress caused by a breach of contract,[20] and, perhaps most importantly, the recalibration of the law relating to liquidated damages clauses and penalties,[21] albeit that a revised (and problematic)[22] version of the 'somewhat formal' distinction[23] between primary and secondary obligations has been retained. In the latter context there has been, for example, renewed emphasis on the proposition that 'the classification of terms for the purpose of the penalty rule depends on the substance of the term and not on its form'[24] and acceptance of the argument that the 'anomalies in the operation of the law as it has traditionally been understood' do not justify abolition of the rule because many of them can be overcome 'by a realistic appraisal of the substance of contractual provisions operating upon breach' and 'by taking a more principled approach to the interests that may properly be protected by the terms of the parties' agreement'.[25]

15 Atiyah and Summers (n 4) 84.

16 [1969] 1 AC 350.

17 [2008] UKHL 48, [2009] 1 AC 61.

18 See further D McLauchlan, 'Remoteness Re-invented?' (2009) 9 Oxford University Commonwealth Law Journal 109.

19 *Sempra Metals Ltd v Inland Revenue Commissioners* [2007] UKHL 34, [2008] AC 561; *Hungerfords v Walker* (1989) HCA 8, (1989) 179 CLR 125.

20 See, eg, *Farley v Skinner* [2001] UKHL 49, [2002] 2 AC 732.

21 *Cavendish* (n 7); *Paciocco v Australia and New Zealand Banking Group Ltd* [2016] HCA 28, (2016) 333 ALR 569; *Wilaci Pty Ltd v Torchlight Fund No 1 LP* [2017] NZCA 152, [2017] NZCCLR 9.

22 See A Summers, 'Unresolved Issues in the Law on Penalties' [2017] Lloyd's Maritime and Commercial Law Quarterly 95, 101–06.

23 *Cavendish* (n 7) [43] (Lord Neuberger and Lord Sumption).

24 Ibid [15] (Lord Neuberger and Lord Sumption).

25 Ibid [39] (Lord Neuberger and Lord Sumption). Interestingly, Lord Hodge, writing extra-judicially, has observed that '[t]he future health of the rule may depend on the court's concentration on substance rather than form and its astuteness to recognise disguised penalties': see Patrick Hodge, 'Revisiting Old Law: Judicial Development of the Law of Contract' in A J Steven, R G Anderson and J MacLeod (eds), *Nothing So Practical as a Good Theory* (Avizandum 2017) 63, 73. However, cf A Summers (n 22) (highlighting, inter alia, uncertainties as to how the 'substance over form' test is to be applied).

However, it would be wrong to suggest that in *Panalpina* Lord Wilberforce was foreshadowing an overarching concern to do justice in the individual case and a wholesale rejection of well-established rules or principles that stand in the way of achieving that. It is true that his Lordship refused to allow technical difficulties in satisfying the formal rules relating to the requirements of offer, acceptance and consideration to stand in the way of enforcing commercial transactions that were clearly intended to be legally binding,[26] but at the same time he believed in the values of freedom and sanctity of contract. In his view, there should be a reluctance to allow excuses for non-performance, particularly in arms-length commercial transactions. Thus, contractual allocations of risk should ordinarily be upheld and parties should not be allowed to escape from contractual obligations merely on the ground of inequality of bargaining power or inadequacy of consideration. Indeed, not long before his judgment in *Panalpina*, his Lordship joined with the other Law Lords in *Photo Production Ltd v Securicor Transport Ltd* in rejecting the doctrine of fundamental breach, and in the course of his speech he said:[27]

> It is significant that Parliament refrained from legislating over the whole field of contract [in the Unfair Contract Terms Act 1977]. After this Act, in commercial matters generally, when the parties are not of unequal bargaining power, and when risks are normally borne by insurance, not only is the case for judicial intervention undemonstrated, but there is everything to be said, and this seems to have been Parliament's intention, for leaving the parties free to apportion the risks as they think fit and for respecting their decisions.

Seven years previously his Lordship had delivered a vigorous dissent in *L Schuler AG v Wickman Machine Tool Sales Ltd*.[28] Unlike the majority, Lord Wilberforce was convinced that the designation of a term in a distributorship agreement as a 'condition' plainly indicated the parties' intention that the contract could be terminated for any breach of the term, however minor. Having reached this conclusion, his Lordship said that:[29]

> . . . to call the clause arbitrary, capricious or fantastic, or to introduce as a test of its validity the ubiquitous reasonable man (I do not know whether he is English or German) is to assume, contrary to the evidence, that both parties to this contract adopted a standard of easygoing tolerance rather than one of aggressive, insistent punctuality and efficiency. This is not an assumption I am prepared to make, nor do I think myself entitled to impose the former standard upon the parties if their words indicate, as they plainly do, the latter.

Although some contract theorists might suggest otherwise, there is no necessary inconsistency between the above statements and Lord Wilberforce's observation in *Panalpina*. It was no doubt his Lordship's view that *ordinarily* the just solution to disputes between commercial parties involves upholding sanctity of contract. The basic task of the court is to seek and give effect to the intention of the parties. Thus, in arms-length commercial transactions justice means holding parties to mutually agreed risk allocations. Justice also means not allowing conceptual or technical problems to defeat the clear intentions of the parties or their legitimate commercial expectations.

The above discussion is intended to provide some context, and support, for my earlier contention that only limited insights are to be gained from seeking an explanation or assessment of modern contract law in terms of whether judicial reasoning tends to be formal or substantive. The picture is much more complicated than that: it varies not only

26 *New Zealand Shipping Co Ltd v A M Satterthwaite & Co Ltd* [1974] 1 NZLR 505 (PC) 510.

27 [1980] AC 827, 843.

28 [1974] AC 235.

29 Ibid 263.

from jurisdiction to jurisdiction, from court to court, from judge to judge, but also within the various branches of the subject. One must also be mindful of the different senses in which the term 'substance over form' is used. In its narrowest sense the term means that the label the parties attach to their agreement is not determinative of its legal categorisation. Thus, an agreement to engage a person as an 'independent contractor' will be classified as a contract of employment if that is its true nature in the light of the rights and duties agreed to. Here, and elsewhere in our commercial law, 'substance' means 'legal substance' so that, in regulating commercial transactions, the courts have historically proceeded on the basis that 'there is no such thing . . . as looking at the substance, apart from looking at the language which the parties have used. It is only by a study of the whole of the language that the substance can be ascertained.'[30] More commonly nowadays, however, the term 'substance over form' is used in the sense that legal forms, rules or technicalities (so-called 'doctrinalism') should not, particularly in the adjudication of private law disputes, be allowed to prevail over 'economic realities' or get in the way of giving effect to the 'merits', the 'justice of the case', or the 'reasonable expectations' of the parties. Indeed, one finds an increasing number of modern cases in which the decision is justified on the basis that any other result would represent 'a triumph of form over substance' in one or other of the latter senses.[31] In my view, this so-called justification is in danger of becoming a ritual incantation in substitution for precise legal analysis and development of the law by reference to exegesis of established precedents or principles.

However, it is more important for the purpose of the theme of this paper to note that there are many aspects of the law of contract where formal and substantive considerations hold sway in varying degrees. In other words, as mentioned earlier, the law is an amalgam, often complex, of both such considerations. Take, for example, the important body of law governing contract interpretation. The 'highly formal'[32] plain meaning rule has been displaced in most common law jurisdictions by the 'less formal and more substantive'[33] contextual approach championed by Lord Hoffmann,[34] but that rule and its close relative, the parol evidence rule, have strongly influenced the survival of another rule that excludes evidence of prior negotiations as an aid to interpretation. Substantive considerations have also been at work. Thus, in *Chartbrook Ltd v Persimmon Homes Ltd*[35] Lord Hoffmann (with whom the other members of the House agreed) held,

30 *McEntire v Crossley Brothers Ltd* [1895] AC 457 (HL) 463 (Lord Herschell). In that case the fact that a conditional sale agreement had the same practical effect as a chattel mortgage provided no basis for treating it as an 'assurance' by the debtor within the meaning of the bills of sale legislation. Similarly, although the absolute assignment of book debts or block discounting of hire purchase agreements have the effect of providing finance to the seller and 'security' for the buyer, their legal substance is ordinarily that of sale: see, eg, *Re George Inglefield Ltd* [1933] Ch 1 (CA); *Chow Yoong Hong v Choong Fah Rubber Manufactory* [1962] AC 209 (PC); and *Lloyds & Scottish Finance Ltd v Cyril Lord Carpets Sales Ltd* (HL, 29 March 1979) reported in [1992] BCLC 609. In the latter case Lord Wilberforce said (at 615–16): 'Commercially, and in its economic result, [block discounting] may not differ from lending money at interest . . . Legally, however, there is no doubt that discounting is not treated as the lending of money and that the asset discounted is not considered as the subject of a charge.'

31 See text following n 76 below.

32 The term used by Atiyah and Summers (n 4) 15, to describe an interpretative method that 'merely focuses on literal meanings of words'.

33 Atiyah and Summers (n 4) 14.

34 *Investors Compensation Scheme Ltd v West Bromwich Building Society* [1998] 1 WLR 896 (HL); *Mannai Investment Co Ltd v Eagle Star Life Assurance Co Ltd* [1997] AC 749 (HL); and *Chartbrook Ltd v Persimmon Homes Ltd* [2009] UKHL 38, [2009] AC 1101. See now *Wood v Capita Insurance Services Ltd* [2017] UKSC 24, [2017] AC 1173.

35 *Chartbrook* (n 34).

after discussing the advantages and disadvantages of the exclusionary rule, that the case for departing from it had not been made out by its critics. He was unconvinced that the often-raised pragmatic objections to doing so – for example, increased uncertainty and cost, potential prejudice to third parties – were outweighed by the advantages. At the same time, however, his Lordship endorsed the two legitimate safety devices of rectification and estoppel by convention which would prevent the rule causing injustice in cases where the evidence established that the parties had reached an agreement or common understanding as to the meaning of the words in question during their negotiations.[36] Therefore, it would be difficult to describe the modern law of contract interpretation as employing formal as opposed to substantive reasoning, or vice versa.[37]

Let us now consider whether the two recent decisions of the UK Supreme Court concerning the assessment of damages for breach of contract provide support for, or require modification of, the arguments in this and the previous part of this article.

3 The *Swynson* case

3.1 The facts

The essential facts for present purposes were as follows. In 2006 Swynson, a company controlled and owned by a Mr Hunt, lent £15m to EMSL to enable a management buyout of an American company, Evo. The loan was made in reliance on a negligently prepared due diligence report by the accountancy firm of Hurst Morrison Thomson (HMT) which subsequently changed its name to Lowick Rose LLP. The report failed to disclose some serious problems concerning Evo's finances, most importantly the insufficiency of its working capital. If the true position had been reported, the transaction would not have gone ahead. In 2007 Evo began to suffer cashflow problems that necessitated a further loan by Swynson to EMSL of £1.75m. However, the company's finances continued to deteriorate and in June 2008 Mr Hunt agreed to make another loan to EMSL through Swynson of £3m. As part of this transaction Mr Hunt obtained an 85 per cent shareholding in EMSL. Nevertheless, despite these cash injections Evo's performance did not improve. In December 2008 Mr Hunt made a fateful decision to refinance the 2006 and 2007 loans. He personally lent over £18m to EMSL on the condition that it be used to repay the latter loans. The loans were duly repaid, and as a result only the June 2008 loan was left owing to Swynson. There were two main reasons for this transaction. First, under UK law once Mr Hunt acquired control of both Swynson and EMSL Swynson

36 In fact, from time to time at least three other mechanisms have been used to give effect to a meaning that was commonly understood or agreed in the course of the parties' negotiations. These are, in brief: first, the 'private dictionary' principle (see, eg, *Partenreederei MS Karen Oltmann v Scarsdale Shipping Co Ltd (The Karen Oltmann)* [1976] 2 Lloyd's Rep 708 (QBD)); second, the 'objective background facts' exception (see, eg, the often endorsed decision in *Macdonald v Longbottom* (1859) 1 E & E 977, 120 ER 1177, where evidence was admitted of a pre-contract conversation showing that a contract for the sale of 'your wool' was intended to include both wool produced on the seller's farm and wool that the seller had bought in from other farms); and, third, Mason J's 'united in rejecting' principle in *Codelfa Construction Pty Ltd v State Rail Authority of New South Wales* (1982) 149 CLR 337 (HCA) 352–53 which, despite being tentatively expressed, has since been widely followed in Australia.

37 Much the same might be said about the law relating to formation and variation of contracts. This is exemplified by the decision of the UK Supreme Court in *Rock Advertising Ltd v MWB Business Exchange Centres Ltd* [2018] UKSC 24, [2019] AC 119. In that case the court buttressed its formal reasons for upholding a 'No Oral Modification' clause with substantive considerations of commercial convenience. Thus, it was said (at [27]) that such clauses serve the legitimate commercial purposes of preventing fabrication of informal variations as a defence to summary judgment applications, avoiding disputes as to the making and terms of such variations, and enabling corporations to more easily police their internal rules concerning authority of employees to bind them.

became liable to pay tax on interest payable by EMSL even though payments were not being made. Secondly, Mr Hunt decided that it was contrary to Swynson's best interests for the impaired loans to remain on its books. No consideration was given to the possibility that the refinancing might jeopardise a future damages claim against HMT.

Despite Mr Hunt's best efforts Evo's fortunes did not improve and eventually the company was wound up. Neither the June 2008 loan nor the refinancing of the earlier loans were repaid, and in 2012 Swynson and Mr Hunt issued proceedings against HMT claiming damages of over £16m (representing the amount of the loans less the sums received from various realisations). HMT did not dispute that it was liable for the amount of the 2008 loan but contended that no damages were recoverable by Swynson in respect of the earlier loans because they had been repaid by EMSL, and accordingly there was no loss suffered. As succinctly explained by Lord Sumption in the Supreme Court, Swynson and Mr Hunt made four arguments in response:[38]

> . . . (i) that the December 2008 refinancing was res inter alios acta and did not affect the amount of Swynson's recoverable loss; (ii) that if the loss was not recoverable by Swynson it was recoverable by Mr Hunt, on the footing that HMT owed him a duty of care; (iii) that Swynson was entitled to recover on the principle of transferred loss; and (iv) that HMT having been unjustly enriched by Mr Hunt's provision of funds to EMSL to repay Swynson, Mr Hunt was subrogated to Swynson's claims against them.

The second argument was rejected by the trial judge and not appealed. The third and fourth arguments found little support in the lower courts[39] and were convincingly rejected in the Supreme Court.[40] A discussion of them is beyond the scope of this article and, in any event, the critical issue in the case in my view concerned the first argument.

3.2 The trial judge's decision

Mrs Justice Rose held that Mr Hunt's refinancing did not affect Swynson's damages claim. It was 'not an act of mitigation which operates to extinguish the loss that Swynson has suffered on the 2006 and 2007 loans for the benefit of HMT'.[41] Applying the principles of mitigation laid down by Viscount Haldane in the *British Westinghouse* case,[42] the judge ruled that the refinancing was 'mere good fortune' and 'not something that Swynson brought about in the ordinary course of business in order to mitigate the consequences of HMT's negligence'.[43] The problem with this reasoning, as later pointed out in the Supreme Court, is that the case raised an issue of avoided loss, not mitigation.[44] In other words, the question was whether the actions of a *third party* had the legal effect of avoiding or reducing the loss claimed, not whether steps taken by the *claimant* in response to the defendant's breach had that effect.

38 *Swynson* UKSC (n 1) [8].

39 Only Sales LJ in the Court of Appeal thought that Mr Hunt had a claim in unjust enrichment if Swynson's damages claim failed. If required, he would have held (*Swynson* EWCA (n 8) [57]) that 'Mr Hunt had made out a good claim in unjust enrichment such that he would have been entitled to be subrogated in equity to Swynson's damages claim against HMT.'

40 But see Peter Watts, 'Lucky Escapes' (2017) 133 Law Quarterly Review 542.

41 [2014] EWHC 2085 (Ch) [54].

42 *British Westinghouse Electric and Manufacturing Co Ltd v Underground Electric Railways Co of London Ltd* [1912] AC 673 (HL).

43 *Swynson* EWHC (n 41) [54].

44 See, eg, *Swynson* UKSC (n 1) [97] (Lord Mance).

3.3 The Court of Appeal

Rose J's decision was upheld by majority in the Court of Appeal (Longmore and Sales LJJ, Davis LJ dissenting). Longmore LJ recognised that the case was strictly one of avoided loss, not mitigation, but nevertheless held that the *British Westinghouse* principles governed both situations. In his Lordship's view, the refinancing arose out of the consequences of HMT's breach 'but in no way did it arise in the ordinary course of business'.[45] He did accept that, ordinarily, where a loan induced by negligent advice is repaid, the repayment must be brought into account if the adviser is sued by the lender, but said that there was 'no inflexible rule to that effect'.[46] He went on to posit the situation where Mr Hunt had made a gift to Swynson of the amount outstanding under the loans to enable the latter to balance its books. There would be no question of that benefit having to be brought into account and a different result should not occur 'merely because the payment [was] made through EMSL'.[47] To hold otherwise 'would be a triumph of form over substance'.[48] Moreover, this did not involve piercing the corporate veil. Rather it was 'merely to disregard technicalities'[49] and to give effect to 'the realities of the case'.[50]

Sales LJ expressed his agreement with Longmore LJ in equally forceful terms. In his view, the principles that determine whether a payment received from a third party after breach is to be treated as 'collateral' or *res inter alios acta* 'are intended to reflect practical reality and basic justice as between the three persons involved: the person who has suffered the loss, the person who is in law responsible for causing the loss and the third party who has made a payment which reduces that loss'.[51] Therefore a court must 'focus on the substance of the matter, as against the technical form which may have been adopted by the third party in choosing how to benefit the person who has suffered the loss'.[52] Acceptance of the approach of the dissenting judge, Davis LJ, 'would place a huge premium on the particular form chosen for a transaction intended to benefit that person, which seems to me remote from the practical reality and justice of the matter'.[53] The fact that Mr Hunt chose for tax reasons to advance funds to EMSL to enable repayment of the loans, rather than directly to Swynson to improve its financial position, was 'not something which affects the substantive position or the just result' as between Swynson, HMT and Mr Hunt.[54] The situation was analogous to the case of benevolent payments where, as explained by Lord Reid in *Parry v Cleaver*,[55] considerations of 'justice, reasonableness and public policy' dictate that the benefit need not be brought into account. Thus, Sales LJ concluded:[56]

> Like benevolent contributors, Mr Hunt certainly did not intend that HMT should benefit from the funding which he was driven to provide to Swynson via EMSL. Nor did he simply pay EMSL money as a general contribution to its fund of assets: the terms of his loan to EMSL were such that it had to pay the money on

45 *Swynson* EWCA (n 8) [17].
46 Ibid [16].
47 Ibid.
48 Ibid.
49 Ibid.
50 Ibid [18].
51 Ibid [53].
52 Ibid [54].
53 Ibid.
54 Ibid.
55 [1970] AC 1, 13 15.
56 *Swynson* EWCA (n 8) [56].

> to Swynson. EMSL had no funds apart from what it received from Mr Hunt with which it could have repaid the 2006 and 2007 loans. In the circumstances, I consider it would be contrary 'to the ordinary man's sense of justice, and therefore contrary to public policy' (to use Lord Reid's language) that the funding Mr Hunt was driven to provide to help Swynson in the difficult position in which it found itself as a result of HMT's negligence should be treated as inuring to the benefit of HMT rather than Swynson alone.

Davis LJ, however, had no doubt that, since the 2006 and 2007 loans were repaid in full, Swynson had suffered no loss in respect of them. The fact that Mr Hunt did not appreciate the effect that his refinancing would have on Swynson's claim against HMT was 'neither here nor there'.[57] His Lordship made, inter alia, the following points in support of this conclusion. First, the majority's reliance on the *British Westinghouse* principles was misplaced because the case was not about whether Swynson had mitigated its loss but whether Mr Hunt's actions had the effect that no loss had been suffered. It was therefore immaterial whether the refinancing was 'in the ordinary course of business'. Secondly, there was no basis for treating the loan repayments as 'res inter alios acta or some kind of collateral transaction' because they had been brought about by the very party, EMSL, which was in default.[58] Thirdly, a ruling in favour of HMT did not 'give rise to a triumph of form over substance' or fail to have regard to 'the realities of the case'.[59] His Lordship said:[60]

> One cannot simply disregard the actual form which the transactions took, with a view to achieving what perhaps may appear (to some) to be a 'just' result. In fact, as I see it, the form here *is* the substance. The commercial form which the 2008 arrangements took—and designedly took—was that Mr Hunt (not, I note incidentally, acting 'benevolently' but through hard-headed commercial considerations) made the loan to EMSL. EMSL then, as required by clause 3.2 of the Loan Agreement of 31 December 2008, used most of that money to discharge its obligations to repay Swynson. It does not seem to me to be legitimate in such circumstances to say that, in effect, it was 'all Mr Hunt'. This is not a mere technicality. One cannot ignore the corporate structures involved. On the contrary, they must be respected. I am not able to agree with Longmore LJ that the contrary approach is 'merely to disregard technicalities'.

Fourthly, the fact that the repayment by EMSL was due to the 'mere good fortune' of Mr Hunt's intervention provided no support for Rose J's decision. Any refinancing by a borrower where the loan was negligently induced is likely to be 'adventitious' from the defendant's perspective.[61]

3.4 The Supreme Court

As foreshadowed, the Supreme Court was unanimous in allowing HMT's appeal. Judgments were delivered by Lord Sumption,[62] Lord Mance and Lord Neuberger.[63] The outcome was inevitable once it became apparent that the judges would insist on adhering to the separate

57 Ibid [32].

58 Ibid [34].

59 Ibid [35].

60 Ibid [36].

61 Ibid [42]. In fact, Rose J was referring (*Swynson* EWHC (n 41) [54]) to Swynson's, not HMT's, good fortune. This was part of her explanation of why the refinancing was not an action in the ordinary course of Swynson's business and hence did not mitigate the loss.

62 With whom Lord Neuberger, Lord Clarke and Lord Hodge agreed.

63 With whom Lord Clarke agreed.

legal personalities of Swynson and Mr Hunt. Indeed, one only had to read the first paragraph of Lord Sumption's lead judgment to know the result. His Lordship said:[64]

> The distinct legal personality of companies has been a fundamental feature of English commercial law for a century and a half, but that has never stopped businessmen from treating their companies as indistinguishable from themselves. Mr Michael Hunt is not the first businessman to make that mistake, and doubtless he will not be the last.

In rejecting the conclusion of the lower courts that the repayment of the 2006 and 2007 loans need not be brought into account, the following main points were made. First, this was a case of avoided loss, not mitigation, and therefore, in the words of Lord Neuberger, the reasoning in the *British Westinghouse* case was 'simply not in point'.[65] Secondly, the repayment of the loans 'discharged the very liability whose existence represented Swynson's loss' and therefore 'could not possibly be regarded as collateral'.[66] The fact that EMSL borrowed the money from Mr Hunt 'was no more relevant than it would have been if it had been borrowed from a bank or obtained from some other unconnected third party', particularly since HMT did not owe a duty of care to Mr Hunt.[67] The repayment could not be treated as discharging the loans 'as between Swynson and EMSL, but not as between Swynson and HMT'.[68] Thirdly, Lord Reid's reasoning in *Parry v Cleaver* did not support Swynson's argument. The ruling there that the question whether the likes of charitable payments, pensions and insurance proceeds are to be brought into account depends on 'justice, reasonableness and public policy' was not to be 'treated by judges as a green light for doing whatever seems fair on the facts of the particular case'.[69] Such benefits were truly collateral in nature because they 'are effectively paid out of [the claimant's] own pocket' (as in the case of insurance) or 'the result of benevolence (whether from the government, a charity, or family and friends)'.[70] They did not resemble the benefit in the present case 'where it is suggested that the court can ignore what is, in its intrinsic nature, a repayment of the loan under and by virtue of which the loss has been incurred'.[71] Fourthly, the fact that the position would have been different if Mr Hunt had given the amount of the loans to Swynson did not mean, as the Court of Appeal majority said, that a ruling in favour of HMT would represent a triumph of form over substance. Such a gratuitous payment was different in nature from what occurred in the present case because 'Mr Hunt's loan to EMSL was intended to and did lead to actual payment off of the first two loans which Swynson had made to EMSL'.[72] The same point was put even more strongly by Lord Neuberger who concluded his discussion of the issue by observing:[73]

64 *Swynson* UKSC (n 1) [1]. See also Lord Mance at [46].

65 Ibid [97]. See also Lord Sumption (at [13]) and Lord Mance (at [46]).

66 Ibid [13] (Lord Sumption).

67 Ibid [12] (Lord Sumption).

68 Ibid.

69 Ibid [98] (Lord Neuberger). See also Lord Sumption (at [11]): 'Justice, reasonableness and public policy are . . . the basis on which the law has arrived at the relevant principles. They are not a licence for discarding those principles and deciding each case on what may be regarded as its broader commercial merits.'

70 Ibid [98] (Lord Neuberger).

71 Ibid [47] (Lord Mance).

72 Ibid [48] (Lord Mance). See also his Lordship's statement (at [49]) that 'there is all the difference between a benevolent act which benefits a claimant (here Swynson) collaterally in an amount equivalent to a loss which it has incurred and satisfaction of the claimant Swynson's loss, by Mr Hunt's funding of EMSL to repay Swynson'.

73 Ibid [100].

> It is true that the money provided in the form of the new loan to EMSL could have been made available to Swynson (or even possibly to EMSL) by Mr Hunt in a way which would not have resulted in Swynson's loss being avoided, but that cannot possibly justify the conclusion that it must therefore be treated as if it had that effect. The fact that a transaction could have been differently arranged does not mean that it must have the same consequences as if it had been differently arranged. As a matter of logic, such a proposition would lead to an impossible situation, and as a matter of experience, it is by no means unusual to encounter cases where a transaction could be structured in two (or more) different ways, each of which would have different consequences—both in law and in commercial reality.

3.5 Form over substance?

It might be considered that the judgments of the Supreme Court in Swynson provide an example of highly formal reasoning: Mr Hunt and Swynson were separate legal persons; the repayment of the 2006 and 2007 loans discharged EMSL's liability to Swynson; and therefore the latter could not be said to have suffered a loss as a result of having been induced by HMT's negligence to make those loans. On the other hand, it cannot be said that the judgments are *formalistic* in the sense that the judges applied rules without reference to their underlying justifications and disregarded considerations of justice or convenience. Their Lordships probably had some sympathy for Mr Hunt's position,[74] but it seems that they had no qualms about denying his company's claim. In particular, there were no good reasons for treating the refinancing as analogous to benevolent payments where the courts have held that considerations of justice and public policy require the benefit to be ignored when applying the compensatory principle. The law was simply too clear. Fundamental legal principles dictated a conclusion that no loss had been suffered. The notion that this represented a triumph of form over substance was rejected.[75] If pressed on this matter, the likely response would have been that this was a hard case that should not be allowed to make bad law. If pressed further they might have pointed to the fact that historically the common law courts have shown no propensity to classify or regulate commercial transactions according to their economic substance or function. What matters is their *legal substance* which can only be derived from construing the agreed terms as a whole.[76]

Nevertheless, the question can fairly be asked whether the law was so clear that Swynson's case was hopeless. Obviously, the trial judge and two experienced Court of Appeal judges thought not. And I have more than anecdotal evidence that many other common law judges would agree with them. In 2016, one month before judgment in *Swynson* was delivered, I addressed a session of the Australasian Appellate Judges Conference in which my co-presenter, a senior judge, strongly defended the Court of Appeal's decision and said that he would be surprised if the Supreme Court reversed it. To my knowledge, only one attendee disagreed with that view. Furthermore, the desire of the Court of Appeal majority to reach what they perceived to be the just result despite the technical obstacles was not without precedent in the recent case law. There have been several modern cases in which the courts have allowed claims or denied defences on the basis that to hold otherwise would represent a triumph of form over substance. One

74 See ibid [90] (Lord Mance).

75 Expressly by Lord Mance (at [48]) and implicitly by the other judges.

76 See text at n 30 above.

notable example, albeit itself contentious, is *Shell UK Ltd v Total UK Ltd*[77] where the Court of Appeal held that a duty of care was owed to a beneficial owner of property by a defendant whose negligent actions damaged the property because 'it would be a triumph of form over substance to deny a remedy to the beneficial owner . . . when the legal owner is a bare trustee for that beneficial owner'. [78] Another example, also contentious, is *Menelaou v Bank of Cyprus plc*[79] where the Supreme Court rejected the defence to the bank's unjust enrichment claim because, inter alia, it 'represent[ed] a triumph of form over substance, or, to use the words of Lord Steyn in *Banque Financière* [1999] 1 AC 221, 227C, "pure formalism"'.[80] How many of us would have confidently predicted that the Supreme Court would disavow such an approach in *Swynson*?

4 *The New Flamenco*

4.1 The facts

Unlike *Swynson*, *The New Flamenco* did raise an issue of mitigation in the strict sense because the broad question was whether an alleged benefit accruing to the claimants that arose from steps taken by them after the defendants' breach of contract must be taken into account in the assessment of damages. The basic facts were relatively straightforward. The defendant charterers repudiated the time charter of a small cruise ship owned by the claimants when the charter still had two years to run. The owners accepted the repudiation and, since there was no possibility of arranging a substitute two-year time charter in the market, they decided to sell the ship. The price obtained was $US23.765m. The owners subsequently sued to recover the net loss of profits they would have earned (some €7.5m) if the charter had not been repudiated. However, since the value of the ship when the charter was due to end would have been a mere $7m as a result of a fall in the charter market, the charterers argued that the difference of $16.765m between this sum and the actual sale price must be brought into account. Accordingly, no damages should be awarded because the sum exceeded the lost profits claimed. The arbitrator agreed, ruling that the 'benefit' should be brought into account since the sale was caused by the breach and was a reasonable step taken to mitigate the damage. On appeal to the High Court,[81] Popplewell J disagreed, but, after being reversed by the Court of Appeal,[82] his Lordship's decision was reinstated and his reasoning endorsed by the Supreme Court.

77 [2010] EWCA 180, [2011] QB 86.

78 Ibid [143] (Waller LJ).

79 [2015] UKSC 66, [2016] AC 176.

80 Ibid [99] (Lord Neuberger). See also *Andrews v Australia and New Zealand Banking Group Ltd* [2012] HCA 30, (2012) 247 CLR 205, [13] (to hold that the doctrine of penalties is confined to obligations triggered by a breach of contract 'would elevate form over substance', but cf *Cavendish* (n 7) [42]); *Homburg Houtimport BV v Agrosin Private Ltd (The Starsin)* [2003] UKHL 12, [2004] 1 AC 715, [34]; *Breakspear v Ackland* [2008] EWHC 220 (Ch), [2009] Ch 32, [124]; *Standard Life Assurance Ltd v Oak Dedicated Ltd* [2008] EWHC 222 (Comm), [2008] 1 CLC 59, [92]; *Mauritius Commercial Bank Ltd v Hestia Holdings Ltd* [2013] EWHC 1328 (Comm), [2013] 2 Lloyd's Rep 121, [19]; *Novoship (UK) Ltd v Mikhaylyuk* [2014] EWCA Civ 908, [2015] QB 499, [92]; and *Super-Max Offshore Holdings v Malhotra* [2017] EWHC 3246 (Comm), [59] and [134].

81 [2014] EWHC 1547 (Comm), [2014] 2 Lloyd's Rep 230.

82 [2015] EWCA Civ 1299, [2016] 1 Lloyd's Rep 383 (Longmore, Christopher Clarke and Sales LJJ).

4.2 The judgment of Popplewell J

The main reason for Popplewell J's decision[83] was that the difference between the price obtained from selling the ship after the repudiation and the value of the ship had the charter been performed 'was not a benefit which was legally caused by the breach'.[84] It was not sufficient that the breach caused a reasonable mitigating step and that the mitigating step resulted in a benefit. There had to be 'a direct causative connection between breach and benefit',[85] and in this case the decision to sell was triggered by the breach but it was 'legally independent'[86] therefrom because the owners had a choice whether or not to realise the capital value of the ship at any time, including during the remaining period of the charter.[87] Whether they did so was a matter for their own 'commercial judgment and involved a commercial risk taken for their own account'.[88] Furthermore, it made no difference that the arbitrator had found that the sale was a reasonable step to take in mitigation of loss. The loss that was mitigated was the loss of net profits from the charter, and the sale achieved this by removing the costs involved in operating or laying up the ship.

His Lordship also held that the same result was 'dictated by the policy grounds' that informed the cases in which it had been held that some post-breach benefits, such as donations from third parties and the proceeds of insurance and pension schemes paid for or contributed to by the claimant, should be ignored.[89] Thus, when the owners first acquired the ship they undertook 'the risk of fluctuations in its capital value' and '[t]o allow the charterers to take the benefit of their decision to sell at what turned out to be an opportune moment in market conditions would be to allow the charterers to appropriate the fruits of the owners' investment in the vessel in a way which would be unfair and unjust'.[90]

4.3 The Court of Appeal[91]

The Court of Appeal's unanimous reversal of Popplewell J's decision turned on the fact that there was no available market for a substitute charter at the time of the repudiation. If a substitute charter had been available, the correct measure of damages would have been the difference between the contract and substitute market rates of hire regardless of the fact that the owners did not actually go into the market. In such a situation subsequent

83 For a fuller account of the judge's reasoning, see D McLauchlan, 'Mitigated Loss or Collateral Benefit?' [2016] Lloyd's Maritime and Commercial Law Quarterly 459, 459–60.

84 *New Flamenco* EWHC (n 81) [65].

85 Ibid [64].

86 Ibid [66].

87 Thus, his Lordship said (at [70]): 'Whilst the charter was on foot, the owners might have sold the vessel subject to charter, provided that they did so on terms which required the new owner to perform the charterparty so that they were not putting it out of their power to perform. They might, for example have entered into a back-to-back charter with the new owner so as to become disponent owners vis-à-vis the charterers, or have arranged a novation to which the charterers and the new owner were party. Alternatively they might have sold the vessel on terms that she be delivered to the new owner following expiry of the charter period. Sale of the vessel was a transaction which could, in principle, have occurred irrespective of the breach . . . But even if it be assumed that the existence of the charter would in practice have formed an impediment to sale before its expiry, a sale was nevertheless a transaction *of a kind* which the owners could have undertaken for their own account irrespective of the breach . . .'

88 Ibid [66].

89 Ibid [73].

90 Ibid.

91 For a fuller account of the court's reasoning, see McLauchlan (n 83) 461–63.

gains or losses resulting from the choice not to seek substitute hire are for the claimant's own account.[92] Thus, where a gain is made, it will be the result of an independent decision not arising out of the breach, and therefore it will be ignored in the assessment of damages. However, since there was no available market, the *British Westinghouse* principles applied and therefore the question was whether the alleged mitigating act arose out of the consequences of the breach and in the ordinary course of business. If the owners had taken advantage of opportunities to trade the vessel, say, by spot charter, the profits earned or losses suffered through such trading would be deducted from or added to, as the case may be, the usual measure of damages where a charterer repudiates,[93] and there was no reason why the position should be different if they decided to sell the ship and obtained a price much higher than the value it would have had if the charter to the defendant had remained on foot. They 'had made a considerable profit from the action they took by way of mitigating what would otherwise have been an undoubted loss' and '[t]hat profit arose from the consequences of the breach and should therefore be brought into account'.[94] Furthermore, their Lordships disagreed with Popplewell J that considerations of fairness and justice required the contrary conclusion. Ignoring the benefit derived from the sale would be inconsistent with the reiterations of the 'fundamental' compensatory principle[95] in *The Golden Victory*[96] and *Bunge SA v Nidera BV*:[97] in other words, it would put the owners in a better financial position than if the contract had been performed. Their Lordships also gave short shrift to Popplewell J's argument that the owners' action was legally independent of the breach because they could have sold the vessel at any time, even during the term of the charter. This was regarded as 'somewhat of a side issue since the central questions must always be whether the actual sale was caused by the breach and was by way of mitigation'.[98]

4.4 The Supreme Court

In a judgment delivered by Lord Clarke[99] the Supreme Court held that the owners were not required to bring into account the benefit derived from the sale because that benefit was not caused by the breach nor the result of 'an act of successful mitigation'.[100] The fall in the value of the ship was 'irrelevant because the owners' interest in the capital value of the vessel had nothing to do with the interest injured by the charterers' repudiation',[101] namely, the loss of income from the remaining two years of the charter. The charterers' repudiation was not the legal cause of the sale because the ship could have been sold at any time, even while the charter remained on foot. Indeed, there was 'no reason to assume' that a sale would have 'followed from the lawful redelivery at the end of the charterparty term, any more than it followed from the premature termination'.[102] Accordingly, when the owners did decide to sell they were 'making a commercial decision at their own risk about the disposal of an interest in the vessel which was no part of the

92 *Koch Marine Inc v D'Amica Societa di Navigazione ARL (The Elena D'Amico)* [1980] 1 Lloyd's Rep 75 (QBD).

93 Namely, the difference between the amount of hire and the cost of earning it.

94 *New Flamenco* EWCA (n 82) [39] (Longmore LJ).

95 Ibid.

96 *Golden Strait Corpn v Nippon Yusen Kubishika Kaisha* [2007] UKHL 12; [2007] 2 AC 353.

97 [2015] UKSC 43; [2015] 2 Lloyd's Rep 469.

98 *New Flamenco* EWCA (n 82) [35] (Longmore LJ).

99 With whom Lord Neuberger, Lord Mance, Lord Sumption and Lord Hodge agreed.

100 *New Flamenco* UKSC (n 2) [34].

101 Ibid [29].

102 Ibid [33].

subject matter of the charterparty and had nothing to do with the charterers'.[103] Remarkably, it was held that the position was the same regardless of whether or not at the time of the repudiation a substitute two-year time charter, shorter charters, or no work at all had been available in the market, an important respect in which, unfortunately, the facts were unclear. As discussed elsewhere,[104] the first two situations are, or at least ought to be, uncontroversial in the sense that the refusal to take account of the sale is readily justifiable on the basis that the existence of alternative charters would have enabled the owners to restore as nearly as possible the same position that they would have been in if the contract had not been breached, so that their choice not to do so would have constituted an independent decision for which they were responsible, for better or worse. Hence it would have been the exercise of this choice, not the charterers' repudiation, that was the legal cause of the ensuing financial consequences that turned out to be beneficial.[105]

However, it is difficult to accept that the charterers' repudiation was not the legal cause of the benefit derived from the sale in the third situation where there was no work available for the ship at the time the contract was terminated. In this situation the reality is that the owners had no choice other than to sell the ship and therefore, as the arbitrator and the Court of Appeal concluded, the sale was a 'necessity'.[106] Here the benefit obtained from the sale is legally caused by the repudiation because the sale would not have occurred but for the repudiation and there is no voluntary conduct on the part of the claimants that breaks the chain of causation.

It thus appears that, in holding that *irrespective of the state of the market* the sale was not caused by the breach and could not be regarded as a relevant act of mitigation, the decisive factor for the Supreme Court was that the ship was a capital asset that could have been sold at any time, including while the charter remained on foot.[107] As a result, the sale 'was incapable of mitigating the loss of the income stream'.[108] In my view, this is a formalistic distinction that should have little bearing on the question whether the financial consequences of the owners' conduct in response to the breach ought, in principle, to be brought into account in assessing their damages. Not only does the proposition that the sale 'was incapable of mitigating the loss of the income stream' appear inconsistent with the view their Lordships earlier endorsed that the benefit need not be of the same kind

103 Ibid [32].

104 D McLauchlan and A Summers, 'Mitigation and Causation of Benefits' [2018] Lloyd's Maritime and Commercial Law Quarterly 171.

105 There is authority for the view that only where there was a substitute *two-year* charter would the benefit from sale of the ship be caused, not by the breach, but by the owners' independent decision not to take advantage of the market. Thus, in *Spar Shipping A/S v Grand China Logistics Holding (Group) Co Ltd* [2015] EWHC 718 (Comm), [2015] 2 Lloyd's Rep 407 [222], Popplewell J held that an available market requires 'a like-for-like replacement'. On this view, which was endorsed by Longmore LJ in the Court of Appeal in *The New Flamenco* (n 82) [29], even '[f]our successive six-month charters are not a like-for-like replacement for a two-year charter'. However, as Andrew Summers and I have argued (n 104) 180, the existence of an available market should not be determined in such a binary manner. The underlying question must always be: what steps would restore the claimants *as nearly as possible* to the position that they would have been in but for the breach? In determining whether shipowners are engaged in speculation following a repudiated charter, the question should be, not whether there was an available market for a perfect substitute, but whether the owners made a choice not to avail themselves of the closest substitute that was available.

106 *New Flamenco* EWCA (n 82) [11].

107 In this respect the court endorsed Popplewell J's point that the owners might have sold the ship subject to the charter by employing any one of a number of legal mechanisms available for this purpose: *New Flamenco* EWHC (n 81) [70]. See above n 87.

108 *New Flamenco* UKSC (n 2) [34].

as the loss claimed,[109] but it fails to address the correct question. That question was, not whether the owners had a choice whether or not to realise the capital value of the ship at any time, but whether, in the circumstances they faced at the time of the termination of the charter, they had a choice to put themselves, as nearly as possible, in the position they would have occupied if the contract had been performed. In other words, did they have a choice whether or not to restore their non-breach position? And if the true position was that there was no work available for the ship and thus the sale was a practical necessity, no such choice existed and the resulting benefit ought to be brought into account. It was no doubt for this reason that, as noted earlier, in the Court of Appeal Longmore LJ regarded the suggestion that the owners could have sold the ship at any time subject to the charter as 'somewhat of a side issue since the central questions must always be whether the actual sale was caused by the breach and was by way of mitigation'.[110] In his Lordship's view, the answer to that question hinged on whether there was an available market for the rehire of the ship. If there was, any loss or gain resulting from the exercise by the owners of a choice not to take advantage of that market is not caused by the breach but by their own independent decision and therefore was for their own account not the account of the defendants. However, if there was not an available market, the sale is not the result of an independent decision and 'there may be no alternative form of mitigation available' other than to sell.[111]

5 Conclusion

Neither of the two decisions of the UK Supreme Court that have been the main subject of this article can be regarded as applying wholly formal reasoning. The judgments in *Swynson* come close to doing so, but in reality their Lordships, while alive to the unfortunate consequences of their decision for Mr Hunt, regarded the law as too clear. It cannot be said that they applied settled rules without giving consideration to the justifications underlying them. They would probably say that any other decision would unsettle the law unduly and potentially be productive of inconvenience and uncertainty in the future. Or they might simply invoke Professor Atiyah's concluding observation in the essay referred to earlier:[112]

> Without necessarily condemning all manifestations of [the trend towards abandoning or limiting formal reasoning in the law of contract] this process seems to me one against which we should be on our guard. Formalism, of course, has a bad name; and to make a decision for formal reasons requires a degree of self-discipline; it requires that we should admit that some other person, some other time, some other place, some other procedure, may be the right one for this decision. But that is often true of decisions in a legal system in which, by and large, everyone can be trusted to play his proper part. *When it is true, we should not be afraid to recognize that to decide a case for formal reasons is to decide it for good reasons.*

The judgment of the Supreme Court in *The New Flamenco*, on the other hand, cannot avoid the charge that it employed, at least in one important respect, formalistic reasoning because, as discussed, it placed reliance on a factor that could not be decisive of the

109 Ibid [30].

110 *New Flamenco* EWCA (n 82) [35].

111 Ibid [29]. As Christopher Clarke LJ pointed out (at [50]), in these circumstances there was no good reason for ignoring 'the benefit of a sale made at the top of a falling market when it is that sale which was both the cause of the benefit and the act of mitigation—a circumstance which precludes it being treated as res inter alios acta'.

112 'Form and Substance in Contract Law' in Atiyah (n 10) 119–20 (emphasis added).

causation issue. Professor Atiyah drew a distinction between *formal* reasoning and *formalistic* reasoning, with the former being 'not per se unjustified or wrong' whereas the latter is.[113] In his view, a court is being formalistic when, for example, it gives 'a formal reason which is not merely bad, but out of place',[114] a criticism that it is arguably not unfair to level at the judgment in *The New Flamenco*. However, it would be misleading to infer that the decision was not influenced by substantive considerations, even though such considerations did not feature in the court's own brief conclusions.[115] This is because earlier in the judgment the court had not simply endorsed the reasoning of Popplewell J in general terms but quoted extensively from it, including the passage in which the latter held that the ruling in favour of the owners could be justified on policy grounds.[116] As noted earlier,[117] his Lordship held that it would be 'unjust and unfair' to allow the charterers to appropriate to themselves the fruits of the owners' original decision to invest in the vessel.[118] When that decision was made they '[took] upon themselves the risk of fluctuations in its capital value . . . and of the financial consequences of a decision as to whether or when to sell',[119] and the charterers should not be allowed to take advantage of what turned out to be a beneficial sale as a basis for denying that loss resulted from their wrongful repudiation. Of course, whether the decision to allow the owners' appeal was actually justified on this basis is open to debate. Based on the several discussions I have had with judges and contract scholars, there are many who favour Popplewell J's point that, regardless of the state of the charter market, the just response was to treat the sale of the ship as a matter for the owners' account and risk so that the capital benefit derived should not be regarded as a relevant compensating advantage that had the effect of extinguishing their damages claim. The benefit was essentially no different than the proceeds of an insurance policy paid for by the claimant. Indeed, they might question whether, if the contract had been performed, the ship would necessarily have been sold upon the expiry of the charter period: the owners might have judged that a rise in the charter market, and consequently an increase in the value of the ship, would soon occur. By contrast, others are firmly of the view that the benefit should be brought into account because the sale was a reasonable step taken to mitigate the damage arising from the charterers' repudiation which, due to the fall in the charter market, resulted in the owners being more than $16m better off than if the contract had remained on foot. Accordingly, applying the fundamental compensatory principle, they suffered no loss, and therefore they should recover nominal damages only. Of course, this just goes to show that Lord Wilberforce's just solution that can be ascribed to reasonable people in the position of the parties[120] may be something on which reasonable minds form different opinions.[121]

113 Ibid 95. See also Atiyah and Summers (n 4) 250 (describing formalistic reasoning as the 'inappropriate use' of formal reasoning).

114 Atiyah and Summers (n 4) 250.

115 *New Flamenco* UKSC (n 2) [29]–[35].

116 Ibid [23].

117 See text at n 90 above.

118 *New Flamenco* EWHC (n 81) [73].

119 Ibid.

120 See text above at n 13.

121 The competing arguments concerning the merits of the owners' claim are discussed further in McLauchlan (n 83).

NILQ 70(2): 241–262

How contractual terms determine fiduciary duties: a two-stage process

DAVID GIBBS-KNELLER

Lecturer in Law, University of East Anglia

AND

DEREK WHAYMAN

Lecturer in Law, Newcastle University[*]

Abstract

The determination of the scope of the fiduciary duty of loyalty, when created by contract, is not a unitary process. It is raised following a multi-factorial enquiry, which considers the nature of the engagement, in a first stage. Here, no single factor is conclusive. It is then, in a separate, second stage, reduced by qualifying contractual terms, which are applied almost strictly logically. This second stage uses the contractual doctrines of interpretation and implication. However, since it is a form of the fiduciary doctrine of authorisation, those contractual doctrines are modified according to fiduciary principles. We argue this follows from the underlying nature of the fiduciary obligation as a way of resolving its internal tensions. While this division has not yet been fully recognised in the cases, the courts have been inching towards it. However, not fully recognising this inevitable division and eliding the two stages has led to defective reasoning and outcomes.

Keywords: fiduciary duties; equity; contractarianism; construction; implication; authorisation.

In this article, we seek to identify the processes in which the equitable duty of loyalty, or, equivalently, the fiduciary duty, responds to any contractual terms that form part of the engagement between the parties. It might be thought that the duty of loyalty is created in a single-stage process. One starts with a blank slate, and then positively implies the duty of loyalty if the nature of the engagement – defined by the terms of the contract – demands it. In this, one moves in a single step from having nothing to something.

This account is too simplistic. However, it is clear from the leading cases that the scope of the duty of loyalty does depend on the terms of the engagement.[1] By scope, we mean to which interests it does or does not apply and any particular acts, within those interests, to which it does or does not apply. There must, therefore, be some role for the contract. We propose that the true framework is that there are two distinct stages in which the contract terms influence the duty of loyalty in very different ways.

First, the duty of loyalty is raised if the parties are in one of the recognised fiduciary relationships or because of the existence of factors – determined from the terms of the contract – that demand single-minded loyalty. Here, no one factor is conclusive, making this stage relatively uncertain. This is the wholly positive process, moving from nothing to something.

* Thanks for the anonymous external and internal reviewers, and also to Prof Duncan Sheehan and Prof Richard Mullender for their helpful remarks. All errors remain our own.

1 Eg *Kelly v Cooper* [1993] AC 205 (PC) 213–4; *Henderson v Merrett Syndicates Ltd (No 1)* [1995] 2 AC 145 (HL) 206.

Conversely, the second stage is a wholly negative process, where the terms of the contract, both express and implied, only reduce the scope of the duty of loyalty. The second stage is a form of fiduciary authorisation (meaning where the principal consents to what would otherwise be a breach of fiduciary duty), differing from the conventional process of subsequent or mid-term authorisation mainly in that its terms are defined before the engagement has commenced and not after. Here, the contractual terms, express and implied, are applied almost strictly logically, making it much more certain, although they are still subject to overriding fiduciary principles.

We argue this follows from the purpose of the duty of loyalty. It is said that the duty of loyalty exists to protect the principal and deter the fiduciary from acts of disloyalty by taking away the advantage or profit.[2] However, there is a competing imperative. The principal is entitled, when fully informed of the consequences of doing so, to relax that rule and authorise what would otherwise be a breach. It has been observed that it may very much be in his or her commercial advantage to do so.[3] The tension between these aims is best resolved through such a two-stage process.

The courts have not explicitly recognised this. Indeed, authorisation has not been subject to much critical analysis in judgments or commentary.[4] Nonetheless, they have been inching towards this position as they seek, case by case and issue by issue, to develop the law in accordance with its underlying principles. Thus, in addition to justifying the two-stage process as a matter of principle, we examine the state of the law, showing how far the courts have come towards this outcome and where there is still some way to go.

Recognising the two-stage process explicitly brings three advantages. First, failure to do so results in an elision of the different principles in each stage (for instance, judges treating second-stage issues as purely contractual) which results in faulty reasoning and incorrect outcomes. Second, it makes the law easier to apply because the character and applicable rules of each stage are better illustrated. Third, one can determine the outcome of stage two relatively easily and with a high degree of certainty. While stage one is still inevitably uncertain, by making it clear that some matters only apply to stage two, stage one becomes a little more certain.

We begin in section 1 by considering the tension in the underlying principles and how it can be resolved via a two-stage process. In section 2, we examine the different role of contract law in each stage. Essentially, contractual doctrines can only be used when they are compatible with fiduciary principles, which means the implication of terms in fact only works well in stage two. We continue in section 3 by showing how other contractual principles are modified to make them compatible with fiduciary law, so they are apt for stage two. Finally, in section 4, we conclude by showing how the framework we advance is easier to apply and how it illuminates errors in judges' reasoning and decisions.

2 Eg *Murad v Al-Saraj* [2005] EWCA Civ 959, [2005] WTLR 1573 [74]; *Boardman v Phipps* [1967] 2 AC 46 (HL); *Regal (Hastings) Ltd v Gulliver* [1967] AC 134, [1942] 1 All ER 378 (HL).

3 *Boulting v Association of Cinematograph, Television and Allied Technicians* [1963] 2 QB 606 (CA) 636, 637.

4 Jennifer Payne, 'Consent' in Peter Birks and Arianna Pretto (eds), *Breach of Trust* (Hart 2002) 297, citing *Spellson v George* [1992] NSWLR 666 (NSWCA) 669. The other major piece on authorisation is Ying Khai Liew and Charles Mitchell, 'Beneficiaries' Consent to Trustees' Unauthorised Actions' in Paul S Davies, Simon Douglas and James Goudkamp (eds), *Defences in Equity* (Hart 2018).

1 The two stages in principle

1.1 One stage or two?

It is clear that fiduciary duties can arise not only in traditional fiduciary relationships (such as agent–principal, trustee–beneficiary, solicitor–client and director–company), but also in 'ad hoc' relationships provided the right factors exist. Consider what the courts have said about raising and scoping the duty of loyalty. Most judgments proceed on the basis that, in ad hoc cases, the terms of the contract are important in some rather nebulous way. The classic exposition is in *Hospital Products Ltd v United States Surgical Corporation*, which illustrates this:

> [I]t is the contractual foundation which is all important because it is the contract that regulates the basic rights and liabilities of the parties. The fiduciary relationship, if it is to exist at all, must accommodate itself to the terms of the contract so that it is consistent with, and conforms to, them. The fiduciary relationship cannot be superimposed upon the contract in such a way as to alter the operation which the contract was intended to have according to its true construction.[5]

This passage has been quoted in the stage two cases of *Northampton Regional Livestock Centre Company Ltd v Cowling*,[6] *Hilton v Barker Booth & Eastwood*[7] and *Global Container Lines Ltd v Bonyad Shipping Co (No 1)*[8] and the stage one cases of *Ranson v Customer Systems plc*[9] and *Fujitsu Services Ltd v IBM United Kingdom Ltd*.[10] *Ranson* was cited with approval by Lord Neuberger MR in the stage two case of *Rossetti Marketing Ltd v Diamond Sofa Co Ltd*.[11] Further vague statements can be seen, such as '[t]he precise scope of [the duty of loyalty] must be moulded according to the nature of the relationship'[12] and the defendant's 'capacity to make decisions . . . is inconsistent [with the existence of a general fiduciary relationship]'.[13]

These *dicta* suggest a compendious single-stage process of identifying a fiduciary duty and its scope by considering a range of factors. In *Hospital Products*, the High Court of Australia identified some of these, such as a relationship of trust of confidence, inequality of bargaining power and the absence of arms-length contracting.[14] The English cases have further suggested an undertaking of assumption of responsibility and entrustment of property, affairs, transactions or interests.[15]

Before considering the reasons for having two different stages, it is helpful to illustrate them with some concrete facts. In *University of Nottingham v Fishel*,[16] a stage one case, the question was to which interests of his employer, if any, Dr Fishel was responsible to as a fiduciary. The scope of his duty was determined by considering the many factors. Against

5 (1984) 156 CLR 41 (HCA) 97.

6 [2014] EWHC 30 (QB) [180] this point not raised on appeal: [2015] EWCA Civ 651, [2016] PNLR 5. The issue was disclosure, a stage two issue.

7 [2005] UKHL 8, [2005] 1 WLR 567 [30]. The issue was whether there was an implied term meaning disclosure was not required.

8 [1998] 1 Lloyd's Rep 528 (QB) 546.

9 [2012] EWCA Civ 841, [2012] IRLR 769 [26].

10 [2014] EWHC 752 (TCC), [2014] 1 CLC 353 [123].

11 [2012] EWCA Civ 1021 [21].

12 *New Zealand Netherlands Society 'Oranje' Inc v Kuys* [1973] 1 WLR 1126 (PC) 1130.

13 *Hospital Products* (n 5) 98.

14 Ibid 69ff.

15 *Peskin v Anderson* [2001] BCC 874 (CA) [34] quoted in *Sinclair Investment Holdings SA v Versailles Trade Finance Ltd* [2007] EWHC 915 (Ch), [2007] 2 All ER (Comm) 993 [79].

16 [2000] ICR 1462 (QB).

the existence of the duty of loyalty stood his employer's encouragement of outside consultancy and that he was merely an employee. In favour was the business-like structure of his academic unit and that he had a key role. The court held that he was subject to the duty of loyalty in managing a team of embryologists for his employer.[17] He should have directed them to work for his employer, and so he incurred liability for using them for his own personal benefit. Nonetheless, he was not under a fiduciary duty for work conducted outside of the UK, since he was not acting for his employer there.

In *Ross River Ltd v Waveley Commercial Ltd*,[18] concerning both stages, the parties were joint venturers so the *prima facie* position was that, in the absence of agency or partnership, there is no fiduciary duty.[19] There was a fiduciary duty because Ross River 'reposed a very high degree of trust' in Waveley Commercial (the stage one issue).[20] The key issue was under what terms payment could be made without it being a breach of fiduciary duty. The answer was that it had to be in accordance with clause 10.5 of the contract, providing that the development profit be paid by the fiduciary to Ross River before it could pay itself except: (i) in respect of proper expenses incurred; or (ii) with the agreement of Ross River. This was held to be consistent with the existence of a fiduciary duty (in stage one),[21] but precluded a breach of fiduciary duty provided Waveley Commercial kept to it (stage two).

The different character of each enquiry is immediately apparent. Stage one is the weighing of a multitude of factors. It looks to the contract because the terms of the contract can make out the factors the court looks to, particularly assumption of responsibility. For instance, Dr Fishel's overall responsibility was defined by a multitude of terms. Conversely, stage two is a crystalline, sharply logical and ordered process.[22] Here, the terms of the contract directly define the exceptions to the duty of loyalty, as in Ross River, and the parties' intention, as expressed in the terms, has a much more direct effect on the duty of loyalty.

1.2 Fundamental principles

Now consider how the different characters of each stage may be justified. We argue that this arrangement is the result of the resolution of a tension between three conventional fundamental principles of fiduciary law. Provided one accepts the validity of these principles, the conclusion follows.[23]

17 Ibid 1494.

18 [2013] EWCA Civ 910.

19 *Crossco No 4 Unlimited v Jolan Ltd* [2011] EWCA Civ 1619, [2012] 2 All ER 754 [88]. See also *Murad v Al-Saraj* [2004] EWHC 1235 (Ch) [325]–[341].

20 *Ross River v Waveley* (n 18) [39].

21 Ibid [40], [94].

22 John Rawls, *A Theory of Justice* (Harvard University Press 2009) 36ff uses the term 'lexically ordered'.

23 They are conventional, but can be controverted. See, eg, John H Langbein, 'Questioning the Trust Law Duty of Loyalty: Sole Interest or Best Interest?' (2005) 114 Yale Law Journal 929; Lionel Smith, 'Deterrence, Prophylaxis and Punishment in Fiduciary Obligations' (2013) 7 Journal of Equity 87; Robert Flannigan, 'The Fiduciary Obligation' (1989) 9 Oxford Journal of Legal Studies 285; Ernest J Weinrib, 'The Fiduciary Obligation' (1975) 25 University of Toronto Law Journal 1; J E Penner, 'Distinguishing Fiduciary, Trust, and Accounting Relationships' (2014) 8 Journal of Equity 202; Lionel Smith, 'The Motive, not the Deed' in Joshua Getzler (ed), *Rationalizing Property, Equity and Trusts: Essays in Honour of Edward Burn* (LexisNexis 2003); Stephen A Smith, 'The Deed, not the Motive' in Paul B Miller and Andrew S Gold (eds), *Contract, Status, and Fiduciary Law* (Oxford University Press 2016); Matthew Conaglen, *Fiduciary Loyalty: Protecting the Due Performance of Non-Fiduciary Duties* (Hart 2011); Matthew Conaglen, 'The Nature and Function of Fiduciary Loyalty' (2005) 121 Law Quarterly Review 452; Rebecca Lee, 'In Search of the Nature and Function of Fiduciary Loyalty: Some Observations on Conaglen's Analysis' (2007) 27 Oxford Journal of Legal Studies 327.

The first fundamental principle of fiduciary law is deterrence or prophylaxis. The duty of loyalty exists to protect the principal by deterring the fiduciary from the temptation to put his or her interests ahead of his or her duty or to make an unauthorised profit.[24] It does this by giving remedies over and above those that exist in the law of contract, namely strict liability to account for profits[25] and susceptibility to rescission.[26] By taking away the fruits of the breach – even if the principal suffers no loss – the fiduciary is deterred from pursuing it.[27]

The second is autonomy. In the mid-term authorisation case of *Boulting v Association of Cinematography Television & Allied Technicians*, it was said that:

> [T]he person entitled to the benefit of the rule may relax it, provided he is of full age and sui juris and fully understands not only what he is doing but also what his legal rights are, and that he is in part surrendering them [If so,] there is no reason why he should not relax the rule, and it may commercially be very much to his advantage to do so.[28]

It is trite that parties make bargains for their mutual benefit. If the principal is sufficiently appraised, considering all the circumstances, including his or her knowledge and sophistication, the protective function of fiduciary law is achieved without the need for the duty of loyalty itself.[29] This would have been the case in *Ross River*, and it is only because Waveley Commercial paid itself outside of the agreed strictures that it was liable. One can also easily imagine that a director, with the knowledge and approval of the company, might offer it a good price for its property or an opportunity. As Chitty LJ said, 'the real evil is not the payment of money, but the secrecy attending it'.[30]

The third is that the protective function must prevail. Fiduciary duties are of course voluntarily assumed – one cannot foist the duty of loyalty on someone.[31] The stage one enquiry looks to whether sufficient responsibility has been assumed by the would-be fiduciary. However, there are aspects of the duty of loyalty that do not respond directly to the parties' intentions and prevail over them. The duty of loyalty has an 'irreducible core' that cannot be excluded no matter how hard the parties try.[32] The core provides mandatory rules, which are imposed in direct opposition to the express terms of a contract, reflecting the protective purpose of the duty of loyalty. This is common theme in equity, also seen in the law of mortgages and unconscionable bargains.[33] This makes fiduciary duties hard to

24 *Bray v Ford* [1896] AC 44 (HL).

25 *Boardman v Phipps* (n 2); *Regal Hastings* (n 2); *Keech v Sandford* (1726) Sel Cas Ch 61, 25 ER 223.

26 *Aberdeen Railway Co v Blaikie Brothers* (1854) 1 Macq 461, [1843–60] All ER Rep 249 (HL Sc).

27 Smith, 'Deterrence, Prophylaxis and Punishment in Fiduciary Obligations' (n 23) considers the difference between deterrence and prophylaxis, but it is not material for present purposes.

28 (n 3) 636, 637. See also *Re Pauling's Settlement Trusts* [1962] 1 WLR 86 (Ch) 108: 'having given his concurrence, [the beneficiary should not be able to sue] provided that he fully understands what he is concurring in'.

29 An argument also made by Matthew Conaglen, 'Fiduciary Duties and Voluntary Undertakings' (2013) 7 Journal of Equity 105, 118. Conaglen of course takes the position that the purpose of fiduciary duties is to ensure the proper performance of non-fiduciary duties: Conaglen, *Fiduciary Loyalty* (n 23). Nonetheless, this argument applies even if one takes the purpose of fiduciary duties to be the conventional one of prophylaxis.

30 *Shipway v Broadwood* [1899] 1 QB 369 (CA) 373.

31 *Hospital Products* (n 5) 97; *Williams v Central Bank of Nigeria* [2014] UKSC 10, [2014] AC 1189 [9]. A fiduciary has power, trust and confidence reposed in him or her; there are no constructive fiduciaries: Lionel Smith, 'Constructive Fiduciaries?' in Peter Birks (ed), *Privacy and Loyalty* (Clarendon 1997).

32 *Armitage v Nurse* [1998] Ch 241 (CA) 253. See also Mark Leeming, 'The Scope of Fiduciary Obligations: How Contract Informs, but does not Determine, the Scope of Fiduciary Obligations' (2009) 3 Journal of Equity 181 for instances where the contract is not predominant.

33 *G & C Kreglinger v New Patagonia Meat and Cold Storage Co Ltd* [1914] AC 25 (HL) ('clogs and fetters' and the mortgagee's equity of redemption); *Credit Lyonnais Bank Nederland NV v Burch* [1997] 1 All ER 144 (CA).

conceptualise as contractual terms implied in fact because in contract law express terms trump implied ones.[34] Moreover, as Smith points out, the ability to modify an obligation does not necessarily mean the modifier comes from the same source.[35]

This can be demonstrated by considering the core requirement of disclosure.[36] Consider an engagement where the fiduciary is entitled to take fees or commissions on management activities. While at common law it is perfectly possible to stipulate 'a reasonable commission' for subcontracting work, the details of that commission must be disclosed in order to avoid liability if a fiduciary duty exists.[37] This is perhaps the paradigm case of how fiduciary law's norms override the intentions of the parties. Unless there is sufficient disclosure, the principal's autonomy is impaired and the prophylactic remedies remain.

Exploring the possible outcomes to such an attempt demonstrates not only the existence of the irreducible core, but gives an early indication of a principled need for two stages. Consider, for example, a term providing that disclosure to the 'principal' is not required and the 'principal' is to take independent advice. First, it could be that the reduction in the duties of the 'fiduciary' are such that the duty of loyalty is wholly displaced (in what we characterise as stage one). In *Chan v Zacharia*, Deane J suggested that '[i]t is conceivable that the effect of the provisions of a particular partnership agreement . . . could be that any fiduciary relationship between the partners is excluded'.[38] Second and alternatively, the courts could still find that there is a fiduciary relationship, and any particular terms attempting to exclude the need for disclosure (or indeed another core facet of the duty of loyalty) would be ineffective. A Canadian judge, Moore J, has said that '[t]he fiduciary duty transcends these terms and it is abhorrent for contractual terms to abrogate that duty'.[39] Once the duty is owed, one cannot derogate from its irreducible core.

That is not to say that all terms will go this way. If the term falls in the middle ground, such that it is neither abhorrent to the duty of loyalty (in stage one), it creates a limited exception, reducing the scope of the duty of loyalty in part (in stage two). A very simple example is a non-secret commission. If the fiduciary stipulates a commission of 10 per cent, this is perfectly legitimate. Taking more is clearly a breach of fiduciary duty, so the duty of loyalty is not displaced altogether.

1.3 Resolving the tension

The inevitable consequence of the combination of the protective function and the irreducible core of the duty of loyalty is that, doctrinally, the process of raising it simply cannot be a sharply logical process of applying crystalline, well-defined rules. Instead, using Rose's terminology, it must be a 'muddy' process of determination.[40] Rather than having a set of rules with results of 'yes' or 'no', there are factors with answers of 'maybe' or 'maybe

34 Eg *Marks and Spencer plc v BNP Paribas Securities Services Trust Co (Jersey) Ltd* [2015] UKSC 72, [2016] AC 742 [18].

35 Lionel D Smith, 'Contract, Consent and Fiduciary Relationships' in Paul B Miller and Andrew S Gold (eds), *Contract, Status, and Fiduciary Law* (Oxford University Press 2016) 123ff. He gives the example of contract modifying tort obligations.

36 It is not a free-standing duty: *Item Software (UK) Ltd v Fassihi* [2004] EWCA Civ 1244, [2004] BCC 994.

37 If the size of the commission is a trade custom and the principal aware of it, this is probably sufficient: *Hurstanger Ltd v Wilson* [2007] EWCA Civ 299, [2007] 1 WLR 2351, [36]ff. See also Peter Watts, *Bowstead and Reynolds on Agency* (20th edn, Sweet & Maxwell 2016) para 6–086; the 'half-secret commission' in the example may attract the remedy of account of profits but not rescission.

38 (1984) 154 CLR 178 (HCA) 196.

39 *Penner v Yorkton Continental Securities Inc* [1996] AWLD 456 (Alberta Court of Queen's Bench) [90].

40 Carol M Rose, 'Crystals and Mud in Property Law' (1988) 40 Stanford Law Review 577.

not'. The discussion of *Fishel* is a good illustration of this.[41] It can also be called a 'multi-factorial' or 'range of factors' approach, terminology used in the doctrines of frustration,[42] illegality,[43] mitigation[44] and the common intention constructive trust.[45] It is thought to be generally applicable where the question is whether there is an assumption of responsibility.[46]

There are two reasons why a multi-factorial approach is required. The first is flexibility. If one adopts crystalline rules to determine questions of determination of responsibility, the risk of being driven to unsatisfactory outcomes by those rules or having excessively complicated rules to avoid such outcomes becomes a likelihood. This was the case in the doctrine of illegality, where, after a series of unsatisfactory cases, a majority of the Supreme Court eventually adopted a multi-factorial approach to avoid this problem in *Patel v Mirza*.[47] Moreover, in the context of frustration, it is said that determining the allocation of risk (which is quite like responsibility) was not simply a matter of express or implied provision, but something that takes in 'less easily defined matters such as "the contemplation of the parties"'.[48]

The second is to uphold a standard. Responsibility is always to a certain standard, and when standards are to be upheld, one sees terms such as 'reasonable care', 'satisfactory quality' or 'good reason'. Holding a person to a standard and not merely a checklist of rules inherently favours muddy rules over crystalline ones. It means holding that person to the spirit as well as the letter of the law, but if this is to be legally enforceable, it is the spirit and not the letter that must be binding. If it were the letter – which it would be if the determination were a series of tests with yes/no answers – there would be a considerable danger of imaginative contracting-out or working-around.[49] Given the vulnerability of the principal, this is all the more important. This muddy rules approach, as Rose points out, is very much in the spirit of Holmes' 'bad man'.[50] Judges have constantly justified fiduciary duties as responding to the realities of 'human nature' or 'human infirmity'.[51] The process of raising the duty of loyalty must therefore take this approach.

Curiously, the same premise, that we must plan for the 'bad man' fiduciary, leads to the different conclusion that stage two must be crystalline. A simple economic and behavioural argument is that, unless parties can be sufficiently sure of their responsibilities, they will not

41 Above, text from n 16.

42 *Edwinton Commercial Corp v Tsavliris Russ (Worldwide Salvage & Towage) Ltd (The Sea Angel)* [2007] EWCA Civ 547, [2007] 2 Lloyd's Rep 517 [111].

43 *Patel v Mirza* [2016] UKSC 42, [2017] AC 467 [83], [107].

44 *LSREF III Wight Ltd v Gateley LLP* [2016] EWCA Civ 359, [2016] PNLR 21 [38].

45 *Jones v Kernott* [2011] UKSC 53, [2012] 1 AC 776 [61].

46 *Transfield Shipping Inc v Mercator Shipping Inc (The Achilleas)* [2008] UKHL 48, [2009] 1 AC 61 [93] (Lady Hale).

47 *Patel v Mirza* (n 43). Cf the very complex crystalline approach, putting exception upon exception, proposed by Lord Sumption in the minority. See Paul S Davies, 'The Illegality Defence: Turning Back the Clock' [2010] Conveyancer and Property Lawyer 282 in addition to *Patel* for further illustrations of the problems.

48 *The Sea Angel* (n 42) [111]. The matter of allocation of risk was key in *Davis Contractors Ltd v Fareham UDC* [1956] AC 696 (HL) 743.

49 Rose (n 40) 592–93; Rawls (n 22) 36ff; H L A Hart, *The Concept of Law* (Oxford University Press 2012) 131ff; Duncan Kennedy, 'Form and Substance in Private Law Adjudication' (1976) 89 Harvard Law Review 1685; George P Fletcher, 'The Right and the Reasonable' (1984) 98 Harvard Law Review 949, 962ff; cf Bruce Chapman, 'Law, Incommensurability, and Conceptually Sequenced Argument' (1997) 146 University of Pennsylvania Law Review 1487. Chapman's counter-proposal requires the use of weights to facilitate a more ordered, crystalline approach, which are suspiciously malleable, like factors.

50 Rose (n 40) 592; Oliver Wendell Holmes, *The Common Law* (Little, Brown & Co 1881).

51 Eg *Bray v Ford* (n 24) 51; *Keech v Sandford* (n 25); *Ex p Bennett* (1805) 10 Ves Jun 381, 394; 32 ER 893, 897.

contract. If there is a risk of a party becoming a fiduciary, that party needs to be sure of what to do so that he or she will still gain a benefit from the engagement without the risk of losing it all via the account of profits remedy. While the 'bad man' cannot be allowed to evade his responsibility, he must be allowed to know where he stands. This is all the more important when the remedy of account of profits might give the principal a windfall in such cases, something the courts are wary of doing.[52] This means crystalline, or at least sufficiently crystalline, rules are required. Then, since stage one cannot have crystalline rules, we are driven to the conclusion that there must be a separate stage, which can.

Nonetheless, that is as far as it goes. It does not follow from this argument that authorising terms, such as clause 10.5, are paramount, even if they are agreed by both parties. The argument does not displace the need for the dominance of fiduciary principles. The most one can say is that the precise terms agreed are applied in a crystalline way in stage two when fiduciary principles make that possible and not otherwise. Granted, this may not fully resolve the tension, but it must greatly reduce it.

This leads to the key question of the role of contractual doctrines in stages one and two. The argument thus far goes only to the character of the rules. It does not mandate the presence or absence of contractual doctrine in either, despite these early indications. There are two strands to explore in answering this question: normative and descriptive. Respectively, we consider the role contract law could have in principle and to what extent the courts have adopted it in practice.

2 The role of contract law

2.1 A compatible role for contract

To determine the role of contractual doctrines in creating and scoping the duty of loyalty, one must determine its role both in its home field – interpretation and the implication of terms in fact in the law of contract – as well as in fiduciary law. This is for two reasons. The first is because it might be thought that, even in a two-stage process, the creation of the duty of loyalty is a form of contractual implication, as Edelman has suggested.[53] The second is because, if this is not the case, implication is indeed used in the cases and this must be explained.[54]

The view that implied terms create and scope the duty of loyalty is, at first blush, attractive, because if the duty of loyalty can be scoped by the terms of the contract, one rather suspects they can create it too. Then, contract law and fiduciary law are similar in purpose, divided only by doctrine but ultimately doing the same things.[55] Moreover, by confirming that ad hoc fiduciary duties can be created from contracts in addition to arising

52 See, eg, *Sinclair Investments (UK) Ltd v Versailles Trade Finance Ltd* [2011] EWCA Civ 347, [2012] Ch 453 [47]; *Murad v Al-Saraj* (n 2) [82].

53 James Edelman, 'When do Fiduciary Duties Arise?' (2010) 126 Law Quarterly Review 302. See also James Edelman, 'The Role of Status in the Law of Obligations' in Andrew S Gold and Paul B Miller (eds), *Philosophical Foundations of Fiduciary Law* (Oxford University Press 2014); John H Langbein, 'The Contractarian Basis of the Law of Trusts' (1995) 105 Yale Law Journal 625, 631: 'the contractarian character of the trust is transparent'; Frank H Easterbrook and Daniel R Fischel, 'Contract and Fiduciary Duty' (1993) 36 Journal of Law and Economics 425; Henry N Butler and Larry E Ribstein, 'Opting out of Fiduciary Duties: A Response to the Anti-Contractarians' (1990) 65 Washington Law Review 1; Robert H Sitkoff, 'The Economic Structure of Fiduciary Law' (2011) 91 Boston University Law Review 1039.

54 *Kelly v Cooper* (n 1); see also *Hilton v BBE* (n 7); *Rossetti* (n 11); *Northampton Livestock* [2014] EWHC 30 (QB).

55 Edelman, 'When do Fiduciary Duties Arise?' (n 53) 303 also notes they deal with the same difficult remedial issues.

from 'status' – being in one of the standard fiduciary relations – *Hospital Products* supports, to some extent, this view.[56]

However, there are two main difficulties. The first is that, because of the irreducible core and the dominance of fiduciary principles, sometimes the implied term will need to prevail over express terms to the contrary in order to create the irreducible core.[57] The second is that a term will not be implied in fact unless it is necessary for the business efficacy of the agreement[58] or it goes without saying,[59] and it is sufficiently certain.[60] Terms will not be implied simply because it is reasonable to do so, even if this would reflect the intentions of the parties.[61] Fiduciary duties are rarely necessary for the contract to function. For instance, the buyer of a hotel still got the hotel despite the purchasing agent taking a secret commission, which merely drove up the price.[62]

There was once a possible route around these problems. For a brief period following *A-G of Belize v Belize Telecom Ltd*,[63] it was thought that the doctrine of terms implied in fact had been expanded considerably such that the necessity requirement had been diluted.[64] In this case, Lord Hoffmann characterised the process of implying a term in fact as a form of contractual interpretation (or construction) where the ultimate purpose of the exercise was to convert the common intentions of the parties into contractual terms and therefore this function could take precedence over merely doctrinal requirements such as necessity.[65] One might also speculate that this expansion may have permitted implied terms to prevail over express ones. Edelman expressly relied on this in making his argument that the duty of loyalty was implied in fact as though a contractual term.[66]

However, any travel in this direction was reversed by the Supreme Court case of *Marks and Spencer plc v BNP Paribas Securities Services Trust Co (Jersey) Ltd*. It was 'wrong in law' to say

56 See above, text to n 5.

57 See above, text to n 34.

58 *The Moorcock* (1889) 14 PD 64 (CA).

59 *Shirlaw v Southern Foundries* (1926), Ltd [1939] 2 KB 206 (CA).

60 *Marks & Spencer v Baird Textiles Holdings Ltd* [2001] EWCA Civ 274, [2002] 1 All ER (Comm) 737.

61 *Liverpool City Council v Irwin* [1977] AC 239 (HL).

62 *FHR European Ventures LLP v Cedar Capital Partners LLC* [2014] UKSC 45, [2015] AC 250.

63 [2009] UKPC 10, [2009] 1 WLR 1988.

64 *F&C Alternative Investments (Holdings) Ltd v Barthelemy (No 2)* [2011] EWHC 1731 (Ch), [2012] Ch 613 [225]: 'There are similarities between the reasoning by which terms may be implied into a contract and the way in which fiduciary obligations may be found to arise in a contractual context, and it may be that with the new, unified approach to the question of implication of contract terms set out in *AG of Belize v Belize Telecom Ltd* [2009] 1 WLR 1988 the law is moving towards some assimilation of the relevant tests (see the discussion in Edelman, "When Do Fiduciary Duties Arise?" (2010) 126 LQR 302), albeit the two processes have traditionally been conceptualised as different.' This passage was quoted in *Ross River Ltd v Waveley Commercial Ltd* [2012] EWHC 2487 (Ch), [243]. For commentary, see Gerard McMeel, 'The Rise of Commercial Construction in Contract Law' [1998] Lloyd's Maritime and Commercial Law Quarterly 382; Sir Christopher Staughton, 'How do Courts Interpret Commercial Contracts' (1999) 58 Cambridge Law Journal 303; Richard Buxton, '"Construction" and Rectification after *Chartbrook*' (2010) 69 Cambridge Law Journal 253; David McLauchlan, 'The Lingering Confusion and Uncertainty in the Law of Contract Interpretation' [2015] Lloyd's Maritime and Commercial Law Quarterly 406. Even those who emphasise the fact that Lord Hoffmann did not introduce new principles of construction accept the change of emphasis, eg, Lord Bingham of Cornhill, 'A New Thing under the Sun? The Interpretation of Contract and the *ICS* Decision' (2008) 12 Edinburgh Law Review 374. See Kim Lewison, *The Interpretation of Contracts* (5th edn, Sweet & Maxwell 2011) para 1.01ff.

65 *Belize Telecom* (n 63) [18]. It is possible to cleave a distinction between construction and interpretation (see J W Courtney and Wayne Carter, '*Belize Telecom*: A Reply to Professor McLauchlan' [2015] Lloyd's Maritime and Commercial Law Quarterly 245, 248ff), but we take them to mean the same thing).

66 Edelman, 'When do Fiduciary Duties Arise?' (n 53) 317.

Belize Telecom had expanded the jurisdiction of the courts to imply terms.[67] The process of interpretation has limits,[68] and a term will not be implied unless it is necessary for the contract to function.[69] The foundation for Edelman's argument has been swept away.

Most importantly, the fall of *Belize Telecom* prompted a reconsideration of what these contractual doctrines are for. Carter and Courtney propose that the contractual construction of terms is internally structured and has three aspects: (1) to apply the requirements of certain contract doctrines; (2) to rebut a presumption of intention; and (3) the interpretative function.[70] Function (1) is extensible to non-contractual doctrines such as fiduciary law. Call that function (1B). Carter and Courtney consider that Lord Hoffmann's approach elevated the interpretative function above the others.[71] Upon its retrenchment, it is back in the right role, feeding into other doctrines rather than supplanting them.[72] We must consider how they might work with fiduciary law, but, first, consider a brief examination of how they work with contractual implication, both to illustrate Carter and Courtney's framework, but also as a preliminary step in assessing the compatibility of contractual implication with fiduciary law.

The third function needs no introduction. We must interpret what the parties mean before their intentions can be applied in any respect. The second is not relevant for present purposes. The first is precisely in point and describes the role for interpretation in doctrines where more than mere interpretation is required. Consider *The Moorcock*. The parties agreed to moor the claimant's ship at the defendant's jetty. The question was whether a term could be implied providing the dock be deep enough. The interpretative function had little work to do here. It was the particular doctrine of implication in fact that was in play:

> [T]he law is raising an implication from the *presumed* intention of the parties with the object of giving, to the transaction such efficacy as both parties must have intended that at all events it should have.[73]

The real work is being done by the legal rules in function (1) – a 'presumed intention' that may not recognise actual intention (even if a common intention) and a requirement that any implied term is necessary for the efficacy of the contract. The interpretative function is filtered by this rule of law and subordinate to it.

Thus, at a high level of abstraction, this conventional theoretical role for interpretation fits the non-contractual doctrine of raising and scoping the duty of loyalty. The terms of the contract, interpreted using the usual contractual approaches, go into the two stages. In the first stage, they are applied in the multi-factorial approach of determining if one party has assumed fiduciary responsibilities. In the second, they are applied to reduce those fiduciary obligations.

Difficulties may occur at a lower level of abstraction. At stage one this is unlikely, for two reasons. First, an implied term is unlikely to make a material difference to the multi-factorial enquiry. It is just one factor among many and, since it was derived from the same source as the other factors, it will point the same way. Second, there is little chance of problems with

67 (n 34) [31], [69].

68 Ibid [29], [31].

69 Ibid [21], [23].

70 J W Carter and Wayne Courtney, 'Unexpressed Intention and Contract Construction' (2017) 37 Oxford Journal of Legal Studies 326, 334.

71 Ibid 355.

72 For another critique of contract being called to do more than it is suitable for in another context, see Margaret Jane Radin, 'The Deformation of Contract in the Information Society' (2017) 37 Oxford Journal of Legal Studies 505.

73 (n 58) 68 (emphasis added).

the logic. With sharply logical rules, each with a yes/no answer, the correctness of the result depends on them all being correct. In the multi-factorial first stage, if the interpretative function or any other rule gives a wrong result, that is just one factor of many and, when weighing them all up, the judge can take that into account and explain it away with a new factor accordingly. Implication is therefore broadly irrelevant to stage one.

Consider, then, how contractual terms would work in stage two. Recall how they cannot displace all fiduciary obligations. We now sharpen our proposal: stage two is the application of the usual contractual doctrines, interpretation and implication, and then any relevant fiduciary principles *afterwards*. This is the simplest process that comports with the underlying principles and takes into account the existence of the contract. It allows the detailed terms of the contract to determine the precise scope of the authorisation, and then fiduciary principles are applied thereafter to ensure it meets fiduciary standards, rejecting them if not.

The complication is that the contractual terms that feed into this stage could be both express and implied. For express terms, there is no difficulty. The express terms will reduce the duty of loyalty provided they do not conflict with the overriding fiduciary principles. Implied terms are more troublesome. If the law uses the contractual doctrine of implied terms, it is relying on both the interpretative function, the particular doctrine of contract law (namely implication) and then a *subsequent* fiduciary doctrine – functions (3), then (1) and then (1B) rather than (3) then (1B). While interpretation is relatively unobtrusive, implication is much more so in that it contains contractual principles and norms, particularly the necessity test, which was not designed to work harmoniously with fiduciary law.

However, it so happens that, at stage two, the same principle of necessity is apt in fiduciary law. In contract law, one function of the necessity requirement is to act as a proxy for evidence of the parties' true intentions in order to allow the court to be confident[74] the implied term was indeed intended.[75] It is relatively safe to take a party to have agreed to something if it is necessary for the engagement to function. In fiduciary law, the requirement is that the fiduciary be restrained from committing a disloyal act. If, however, that act is necessary for the engagement to function, then it can hardly be considered disloyal. It may be that 'necessity' must be construed more restrictively, but, in broad terms, it is the right basis. The flip-side is that, because the test is notionally the same, it is easy to forget that its purpose is not, and additional steps may be needed to fulfil that purpose.

Rather surprisingly then, the conventional doctrine of terms implied in fact turns out to be compatible with fiduciary law. At stage one, it does nothing of substance. But at stage two, it is suitable for use as one of that stage's sharply logical rules.

2.2 Authority for the role of contract

Nonetheless, having established their compatibility, we must now establish a positive case for the reception of contractual interpretation and implication as the input to an ultimately fiduciary process in stage two. Otherwise, the argument that stage two should be wholly non-contractual would stand unopposed and our framework would be of purely academic interest. Certainly, mid-term authorisation does not necessarily have to be contractual; it can even be by conduct alone.[76]

There are indeed a significant number of authorities where it has. This supports our claim that the courts are edging towards the two-stage framework. Consider first the

74 A term used in *Reigate v Union Manufacturing Co (Ramsbottom), Ltd v Elton Cop Dyeing Company Ltd* [1918] 1 KB 592 (CA) 603.

75 *Peng v Mai* [2012] SGCA 55, [2012] 4 SLR 1267 [35].

76 *Queensland Mines Ltd v Hudson* (1978) 18 ALR 1 (PC).

proposition that stage two is indeed a form of authorisation. This is essential if one is to be able to draw from the wider body of authorisation case law to see how contractual authorisation would work and indeed to rebut the argument that stage two is purely contractual.

There is apparently no direct authority to this effect. Most cases are concerned with the quotidian task of determining whether the disclosure had been sufficient, sometimes at great length.[77] However, some authorities do indeed see authorisation as an umbrella doctrine that takes different forms. In *Lewin on Trusts*, it is claimed that the rules of authorisation for making a profit from one's office are the same as those for trustees dealing with trust property for their own benefit.[78] In *Knight v Frost*, Hart J applied *dicta* from *Re Pauling's Settlement Trusts* that, following analysis of the nineteenth-century trusts cases, concluded that the same test of consent applies to company cases where the fiduciary is a director.[79] These are all mid-term cases. However, the underlying instrument may expressly or impliedly authorise a conflict in settlement trusts[80] and will trusts,[81] which are not. It goes little further to say the underlying authorising instrument could be the contract creating the fiduciary duty too. More to the point, the outcome of inductive reasoning is clear: the simplest way of reconciling these authorities is to see authorisation as a general doctrine, applying to all fiduciaries, whether at the outset or mid-term.

Consider now support for the proposition that contractual doctrines then feed into the fiduciary doctrines in the (3)/(1)/(1B) sequence. In *Bank of Credit and Commerce International SA v Ali* (concerning a release from contractual liability), it was said that there is no separate equitable doctrine of construction.[82] While the absoluteness of this proposition can be doubted (and on this, see section 3), the differences appear to be only contextual and the *fundamentals* of interpretation are the same. The contractual principles of interpretation have been applied in the context of certainty of intention of trusts,[83] to interpret provisions as to the identity of the subject matter[84] and to determine the rules of pension schemes (which are trusts).[85] The scope of the authorisation, again at least for trusts, is a matter of construction.[86] Finally, in *Sargeant v National Westminster Bank plc* the court indeed went on to apply fiduciary principles after construing the relevant clauses.[87]

Consider now implied terms. They tend to appear in cases where a fiduciary acts for multiple principals where the permission to do so is not spelled out expressly. Then, it is a difficult and contentious issue as to whether such a term is implied. In *Kelly v Cooper*, the issue was that an estate agent (and fiduciary) sold two houses, belonging to different sellers (and principals), to the same buyer. The houses were adjacent and the obvious inference was that the buyer was particularly interested in them for that reason. One can infer that a higher

77 Eg *Re Pauling* (n 28).

78 Lynton Tucker, Nicholas Le Poidevin and James Brightwell, *Lewin on Trusts* (19th edn, Sweet & Maxwell 2015) para 20-095.

79 [1999] BCC 819 (Ch) 828; *Re Pauling* (n 28) 108.

80 *Wright v Morgan* [1926] AC 788 (PC) 796.

81 *Re Beatty* [1990] 1 WLR 1503 (Ch) 1506; *Wright v Morgan* (n 80).

82 [2001] UKHL 8, [2002] 1 AC 251 [17], [25], [44], [79].

83 *Staden v Jones* [2008] EWCA Civ 936, [2008] 2 FLR 1931 [18]ff; see also *Hageman v Holmes* [2009] EWHC 50 (Ch), [2009] 1 P & CR DG17.

84 *Brudenell-Bruce v Moore* [2012] EWHC 1024 (Ch), [2012] WTLR 931.

85 *Premier Foods Group Services Ltd v RHM Pension Trust Ltd* [2012] EWHC 447 (Ch), [2012] Pens LR 151; *PNPF Trust Co v Taylor* [2010] EWHC 1573 (Ch), [2010] Pens LR 261.

86 *Sargeant v National Westminster Bank plc* (1990) 61 P&CR 518 (CA) 521; *Edge v Pensions Ombudsman* [2000] Ch 602 (CA).

87 (n 86).

price might have been obtained as a consequence. The claimant seller's arguments were that non-disclosure of this information was a breach of fiduciary duty, as was the conflict of interest in obtaining a surer commission on both sales over the risk of losing one or both in the process of negotiating and trying to drive the prices up.

The Privy Council rejected these arguments, holding that 'like every other contract, the rights and duties of the principal and agent are dependent upon the terms of the contract between them, whether express or implied'.[88] The necessity requirement applied: 'despite this conflict of interest, [residential] estate agents must be free to act for several competing principals otherwise they will be unable to perform their function'.[89] A term permitting it was implied accordingly. Similarly, in *Hilton v Barker Booth & Eastwood*, Lord Walker made express reference to the traditional tests for implied terms.[90]

The necessity requirement makes the availability of such terms very limited. Such a term could also apply to investment intermediaries, who often act for multiple principals.[91] The fiduciary would still be constrained by any other duties; the authorisation is narrow in scope.[92] Permission to act for multiple principals is as far as it goes.[93] Such a term would not permit the fiduciary to act in a manner that conflicts with his or her principal's interests[94] or for multiple principals in the same transaction.[95] Moreover, in *Hilton*, Lord Walker said that the proposition that one breach of duty owed to one principal could exonerate a breach of a duty owed to another 'seems contrary to common sense and justice'.[96] In *Rossetti Marketing Ltd v Diamond Sofa Company Ltd*, Lord Neuberger MR noted that, whilst a residential estate agent must be free to act for multiple principals, the same could not be said for an agent who sold furniture.[97] Most recently, in *Northampton Regional Livestock Centre Company Ltd v Cowling*, the judge noted that an implied term permitting the individual to act for conflicting principals was only available in situations where such multiple undertakings were inherent to the business.[98]

In these cases, while one sees examples of how the necessity requirement in the doctrine of implied terms is aligned with the necessity requirement of fiduciary law, one also sees unease about going too far. At stage two, even for contractually created fiduciary duties, fiduciary principles are still paramount. This suggests a narrower interpretation of necessity may be appropriate for fiduciary authorisation, where the necessity goes to the necessity of the authorisation rather than merely the business efficacy of the arrangement.

88 *Kelly v Cooper* (n 1) 213.

89 Ibid 214.

90 (n 7) [37].

91 Law Commission, *Fiduciary Duties of Investment Intermediaries* (Consultation Paper) (Law Com CP 215, 2013); *Eckerle v Wickeder Westfalenstahl GmbH* [2013] EWHC 68 (Ch), [2014] Ch 196; Law Commission, *Fiduciary Duties and Regulatory Rules* (Law Com No 236, 1995) para 3.32.

92 See, eg, *Plus Group Ltd v Pyke* [2002] EWCA Civ 370, [2003] BCC 332 [83] citing J Gower, *Principles of Modern Company Law* (6th edn, Sweet & Maxwell 1997) 622; see also *Farrington v Rowe McBride* [1985] 1 NZLR 83 (NZCA) 90. The independence of such duties is discussed in Charles Mitchell, 'Good Faith, Self-Denial and Mandatory Trustee Duties' [2018] Trust Law International 92; see also Companies Act 2006, s 170ff.

93 Acknowledged in *Gwembe Valley Development Co Ltd v Koshy (No 3)* [2003] EWCA Civ 1048, [2004] 1 BCLC 131, [56].

94 See, for example, Law Commission, LC236 (n 91) para 3.31.

95 *North & South Trust Co v Berkeley* [1971] 1 WLR 470 (QB) 482 distinguished in *Kelly v Cooper* (n 1) 215.

96 (n 7) [38].

97 (n 11) [27].

98 (n 6) [186].

3 Four key rules of contractual authorisation

Stage two should therefore be called contractual authorisation.[99] In this section, we summarise and expand the main rules of contractual authorisation, drawing on and expanding the rules in the case law for contractual and mid-term authorisation and taking into account the similarities and differences. We identify four key rules. While not an exhaustive list, these are the most important because of the frequency at which they come up and because they interact with contract law doctrines. This shows the extent to which the framework has gained a foothold in the authorities and the pitfalls to avoid.

3.1 Sufficiency of disclosure and autonomy

It bears repeating: the first key rule of contractual authorisation is that the intention of the parties, ascertained through the conventional contractual interpretation and implication processes, is only accepted as effective authorisation if there is sufficient disclosure. As noted above, this requires the provision of enough information such that the principal can make a fully informed decision. Then, the principal's autonomy is respected and the protective function of fiduciary law is achieved without the need for the duty of loyalty.[100] The courts have been concerned with fleshing out what is sufficient, which is summarised here. There is also one key difference between mid-term and contractual authorisation that warrants exploring: in mid-term authorisation, the engagement is already on foot.

In general, and in accordance with the rigour of fiduciary law, the courts have been strict in their requirements in the mid-term cases. There must be 'full and frank disclosure of all material facts'.[101] The principal 'must be honestly acquainted with all the material circumstances of the case'[102] and 'fully understand[] what he is concurring in'.[103] Not only the existence of the interest, but also its extent, must be disclosed.[104] There must be no *suppressio veri* or *suggestio falsi*, the consent must be clear and must not be obtained through pressure.[105] For company directors, disclosure must be to the board,[106] but also to those independent of the transaction if necessary. Disclosure to a boardroom dominated by those with interests in the transaction is insufficient, and then disclosure to shareholders would be necessary.[107] It has also been said any information that may affect the decision to proceed must be disclosed.[108]

The quotations hint that the requirement is subjective, not objective, meaning that the particular principal must have been sufficiently appraised. It is not enough that a hypothetical 'reasonable principal' in his or her position would have been. For example, a

99 It is tempting to call it authorisation *ex ante* to go alongside (non-contractual) authorisation *ex post*. But consider the case of directors' releases from liability where a second contract is signed. The release is *ex ante* the *second* (releasing) contract but *ex post* the first. This terminology is thus apt to confuse and should be avoided.

100 Above, text to n 29.

101 *Kuys* (n 12) 1227. See also *Re Pauling* (n 28) 107; *Boardman v Phipps* [1964] 1 WLR 993 (Ch) 1011ff affd *Boardman v Phipps* (n 2); *Dunne v English* (1874) 18 Eq 524 (MR) 533.

102 *Lewin* (n 78) para 20-140ff.

103 *Re Pauling* (n 28) 108. See also Payne (n 4) 300.

104 Eg *Gwembe Valley* (n 93) [65]; *Medsted Associates Ltd v Canaccord Genuity Wealth (International) Ltd* [2017] EWHC 1815 (Comm). See John McGhee, *Snell's Equity* (33rd edn, Sweet & Maxwell 2015) para 7-015.

105 *Lewin* (n 78) para 20-140.

106 *Queensland Mines* (n 76).

107 *Gluckstein v Barne*s [1900] AC 240 (HL) 258; *Regal Hastings* (n 2) 150.

108 *Brickenden v London Loan & Savings Co* [1934] 3 DLR 465 (PC) 469; *Johnson v EBS Pensioner Trustees Ltd* [2002] EWCA Civ 164, [2002] Lloyd's Rep PN 309 [69], [72], [83]; see also *FHR European Ventures LLP v Makarious* [2011] EWHC 2308 (Ch), [2012] 2 BCLC 39 [79]; *FHR v Cedar* (n 62).

retail investor would be less well-versed in the risks associated with a particular investment than an institutional investor and would have to be given more information. The *dicta* in *Boulting* also suggest the test is subjective[109] and, in *Farah Constructions Pty Ltd v Say-Dee Pty Ltd*, the High Court of Australia placed great weight on these considerations.[110]

Kelly v Cooper is further authority for the subjectivity requirement. That case was decided as it was because 'their Lordships [were] of the view that . . . the plaintiff was *well aware* that the defendants would be acting also for other vendors'.[111] But some cases are less emphatic. In *Rossetti*, Lord Neuberger said that 'residential estate agents could not sensibly carry out their function if the normal conflict rule applied, and any person instructing an estate agent *must* appreciate that fact'.[112] This could be construed as laying down an objective requirement, as not distinguishing between an objective and subjective requirement, or as a by-the-by observation. In *Hurstanger Ltd v Wilson* the distinction does not appear to have been in point, and the judgment does not consider it.[113]

Since contractual authorisation is a form of the general doctrine of authorisation, these requirements still apply. It is important not to forget them and take the contractual terms as conclusive. This will be in issue where fiduciaries attempt to exclude liability by stipulating that they aim their explanations at a more sophisticated level or, more bluntly, that only limited or no disclosure at all is required.

Consider now the difference in timing. Greater disclosure will often be required for mid-term authorisation outside the original contract than authorisation in that contract. This is a consequence of the fact that the equitable duty of loyalty has a more relational quality than the law of contract.[114] Unlike in a simple transactional agreement, in a long-term engagement the timing of any such agreement and disclosure matters. As time goes on, the information asymmetry problem will worsen as the fiduciary gains more information that might not be passed on to the principal. Moreover, after performance of the engagement has begun and money has been sunk into it, there will be greater pressure to cooperate in order to keep the agreement alive and avoid wasting those costs.[115] Conversely, before the agreement is concluded, the principal still has the opportunity to bargain for better terms, demand more information, consider other business partners and walk away if unsatisfied with the deal on offer.

As an illustration, consider again the case of *Ross River*, where clause 10.5 spelled out the narrow circumstances in which Waveley Commercial could pay itself. The Court of Appeal held this term was sufficient to reduce the duty of loyalty within its narrow parameters. Now consider what the outcome would be if there had been no such term and instead Waveley Commercial had invited its principal to authorise such a payment after the conclusion of the contract. Unless they had been suitably informed of how commercially advantageous this was to Waveley Commercial, sufficient disclosure may not have been made and payment, even within the specified parameters, may have been a breach of fiduciary duty.

109 Above, text to n 28.

110 [2007] HCA 22, (2007) 230 CLR 89 [107].

111 (n 1) 215 (emphasis added).

112 (n 11) [25] (emphasis added).

113 (n 37) [34]ff.

114 T T Arvind, 'Contract, Transactions and Equity' in L A Di Matteo et al (eds), *Commercial Contract Law: Transatlantic Perspectives* (Cambridge University Press 2013); Scott FitzGibbon, 'Fiduciary Relationships are not Contracts' (1999) 82 Marquette Law Review 303.

115 See, eg, Richard A Posner, *Economic Analysis of Law* (9th edn, Aspen 2011) 9 for an economic analysis of 'sunk costs'.

3.2 Burden of proof

The second key rule of contractual authorisation concerns the burden of proof. At the first stage, the principal must adduce evidence that makes out the factors that lead to the imposition of the duty of loyalty. That stage is even-handed in the sense that no presumptions are made against the defendant because it is not yet established that he or she is a fiduciary. While there are apparently no direct statements to this effect in the authorities, there is a conspicuous absence of presumptions against the would-be fiduciary, and there is even-handed consideration of the relevant factors.[116] More positive support can be derived from the courts' attitude that the duty of loyalty should not be imposed instrumentally; in *A-G v Blake Jonathan Cape Ltd (Third Party)* the Court of Appeal said that '[f]iduciary duties should not be superimposed on those common law duties simply to improve the nature or extent of the remedy'.[117]

At the second stage, however, the burden of proof shifts. This is because of the evidential uncertainties and the inherently stronger position of the fiduciary. It is the only way to respect the autonomy of the principal given the information asymmetries between fiduciary and principal and thus uphold the aim of fiduciary law. For mid-term authorisation, this matter has been long settled.[118] For contractual authorisation, the reverse burden of proof was made explicit in *Ross River Ltd v Waveley Commercial Ltd.*[119] Lloyd LJ, with whom Fulford and Mummery LJJ agreed, held that:

> [I]f a fiduciary duty exists at all, it throws the burden on the party subject to the duty to justify any payment in any case where there is any doubt as to whether it was properly made.[120]

This is another instance where the norms of fiduciary law coincidentally match those of contract law; in contract law it is for the propounder of the implied term to prove it.[121] More to the point, however, is how the shifting burden precisely reflects the two different stages.

3.3 Doubts resolved against the fiduciary

It further follows that only a conservative form of interpretation will respect the principal's autonomy. The third key rule is then that any doubts or uncertainties in the agreement as to the scope of the fiduciary duty must be resolved against the fiduciary. This is the contractual principle of *contra proferentem*, where uncertainty is resolved against the profferer. On the facts of the fiduciary–principal relationship where the fiduciary is accused of a breach of fiduciary duty, this uncertainty will always be resolved against the fiduciary.

There are apparently no mid-term authorisation cases where this has been applied terribly clearly, but a few old cases could be interpreted to support the application of *contra proferentem*.[122] Where there was undue influence, the court would not rely on concurrence

116 Eg *Hospital Products* (n 5); Kuys (n 12); *Barthelemy* (n 64) [221]ff; *Fishel* (n 16) 1489ff; *Ranson* (n 9); *Helmet Systems Ltd v Tunnard* [2006] EWCA Civ 1735, [2007] FSR 16; *Peskin v Anderson* (n 15) [34].

117 [1998] Ch 439 (CA) 453 quoting *Norberg v Wynrib* (1992) 92 DLR (4th) 449 (SCC) [147] (Sopinka J), noted in *Fishel* (n 16) 1490.

118 *Tito v Waddell* (No 2) [1977] Ch 106 (Ch) 225 (dealing with trust property); *Rossetti* (n 11) [22]; *Northampton Regional Livestock Company Ltd v Cowling* [2014] EWHC 30 (QB) [188] (affirming *Rossetti*) revd on other matters: [2015] EWCA Civ 651, [2016] PNLR 5 (acting for two principals); *Cobbetts LLP v Hodge* [2009] EWHC 786 (Ch) [108] (fiduciaries); *Hurstanger v Wilson* (n 37) [35] (agents).

119 (n 18). This principle from *Ross River* was applied in *Psycare Ltd v Mundy* [2013] EWHC 4573 [30]ff.

120 *Ross River v Waveley* (n 18) [64]. See also at [95].

121 Carter and Courtney (n 70) 339.

122 *Lewin* (n 78) para 20-140.

'fished out from a loose expression in a letter'.[123] Where the circumstances were suspicious, only the clearest evidence would do.[124] However, in cases of authorisation from the start the courts have expressly applied *contra proferentem*. This has been where clauses in the trust deed have attempted to exonerate trustees or limit their liability for breach of trust[125] and breach of fiduciary duty.[126]

There are two issues to consider: the survival of *contra proferentem* and its compatibility. It is clear that the *contra proferentem* rule is in something of a decline in the law of contract.[127] However, it does appear to have a future, albeit a limited one. In the recent Court of Appeal case of *Persimmon Homes Ltd v Ove Arup and Partners Ltd*, it was said that it still has a role where there is a power imbalance.[128] Fiduciary relations are probably the paradigm case of a power imbalance, and so the continuing existence of *contra proferentem* seems secure here.

As for compatibility, the rule is already context-sensitive. This was observed in *Bogg v Raper*, where Millett LJ noted that it had two aspects. *Contra proferentem* requires one both to construe against the party relying on the clause, and also against the author of the instrument. For contracts, the two aspects are aligned, but for trusts this is not usually the case.[129] There is no reason, therefore, that *contra proferentem* cannot be retained and adapted so it is compatible in stage two.

Two adaptations are required. First, since authorisation looks to the subjective, any terms should be construed with this in mind. The interpretation must be within the range of reasonable interpretations of the authorisation this particular principal actually decided to grant. It must take into account the information asymmetry problem – the principal will know less of the background information, and the words must be interpreted accordingly.

Second, there is the relational nature of a fiduciary engagement, which endures longer than most transactional contracts. One must construe any reducing clause narrowly, with a view to the many possible futures of the engagement and permit the fiduciary the minimum possible. This is best illustrated with the case of the company director. Directors usually have uncircumscribed fiduciary duties because they take responsibility for all possible interests of the company with no specific direction on how the outcome should be achieved.[130] The scope of the duty is set at the time of the initial engagement, namely the appointment of the director, yet any change in circumstances will not be known until later. It would be impossible to start with a fine-tuned duty of loyalty that covers such future events unless they are expressly and very clearly defined.

3.4 Honesty, good faith and negligence

The fourth key rule is that any fiduciary duty-reducing terms are subject to the requirements of honesty and good faith, not something generally found in the law of contract. It is clear

123 *Carpenter v Heriot* (1759) 1 Eden 338, 341; 28 ER 715, 716.

124 *Morse v Royal* (1806) 12 Ves Jun 355, 373; 33 ER 134, 141.

125 *Bogg v Raper*, *The Times*, April 22 1998 (CA) [28]ff; this point affd *Wight v Olswang* (No 1) (1999) 1 ITELR 783 (CA). See generally *Armitage v Nurse* (n 32); *Walker v Stones* [2001] QB 902 (CA).

126 *Barnsley v Noble* [2016] EWCA Civ 799, [2017] Ch 191 (self-dealing). See also the discussion of *Global v Bonyad* (n 8) in section 4.

127 Edwin Peel, '*Whither* Contra Proferentem?' in Andrew Burrows and Edwin Peel (eds), *Contract Terms* (Oxford University Press 2007); Edwin Peel, 'Contra Proferentem Revisited' (2017) 133 Law Quarterly Review 6.

128 [2017] EWCA Civ 373, [2017] PNLR 29 [52].

129 (n 125) [28]ff.

130 *Re Allied Business & Financial Consultants Ltd* [2009] EWCA Civ 751, [2009] BCC 822 [71]; David Gibbs, 'The Absolute Limit of Directors' Fiduciary Liability for Conflicts of Interest: the Director's Perspective' (2015) 36 Company Law 231.

from the case law that these requirements can persist even where self-interest has been permitted.[131] This means the authorisation would not cover acts that require the exercise of discretion unless that discretion is exercised honestly and without self-interest. While Millett LJ said '[a] servant who loyally does his incompetent best for his master is not unfaithful',[132] the flip-side is that a fiduciary who deliberately acts against his or her principal very much is.

The case law on this matter is relatively rare. Since it requires a discretion, it comes up most often in trusts cases. One non-trusts case, regarding disclosure, is *Industrial Development Consultants Ltd v Cooley*, where the director's release from liability was vitiated because he had dishonestly represented that he was in ill health.[133] It was a trusts case, *Armitage v Nurse*, where the general rule was laid down. The trustee may be excused from all liability except that caused by his or her fraud or dishonesty. Millett LJ held that dishonesty:

> . . . connotes at the minimum an intention on the part of the trustee to pursue a particular course of action, either knowing that it is contrary to the interests of the beneficiaries or being recklessly indifferent whether it is contrary to their interests or not.[134]

While this test is commonly applied to the selection of investments,[135] it could well be applied to a fiduciary joint venture. In the absence of a suitable reported case, consider this hypothetical. Suppose there is a term providing that a fiduciary joint venturer 'is wholly entitled to "highly speculative" drilling opportunities where, in its opinion and such opinion is final, the chance of striking oil is less than 50 per cent'. The fiduciary would only have to honestly believe the chance was less than 50 per cent under that definition. The subjective 'own opinion' provision would displace the contractual duty to evaluate such opportunities with reasonable care and skill (an objective standard).

This has given rise to much criticism, including judicial dissent.[136] For professional trustees the test was modified in *Walker v Stones*, but only so far, such that, if no reasonable trustee would have considered their actions in the interests of the beneficiaries, this too would be outside the exemption clause.[137] This makes the test objective. This would mean that the hypothetical exclusion clause would not be effective. It follows, perhaps surprisingly, from *Walker v Stones*, that a non-excludable duty to act with reasonable care and skill – a true negligence standard – is fixed upon professional trustees, and by extension, professional fiduciaries. One may quibble about whether a certain class of person is 'professional' for these purposes, but, if *Walker v Stones* applies, that is the result.

At first blush, this seems wrong. Investment and management decisions are ordinarily not fiduciary in character. They ought, therefore, to be subject to non-fiduciary norms, and thus liability for breach ought to be subject to exclusion except when the breach is dishonest (and therefore takes on a fiduciary character). The surprising outcome is that in the authorisation stage a negligent fiduciary decision may well be a breach of fiduciary duty no matter what any authorising terms provide for. Nonetheless, given the controversy over this standard, judicial revision of it would not be unexpected.

131 Mitchell (n 92).

132 *Bristol and West Building Society v Mothew* [1998] Ch 1 (CA) 18.

133 [1972] 1 WLR 443 (Birmingham Assize).

134 (n 32) 251.

135 Eg *Walker v Stones* (n 125); *Fattal v Walbrook Trustees (Jersey) Ltd* [2010] EWHC 2767 (Ch), [2012] Bus LR D7; *Spread Trustee Co Ltd v Hutcheson* [2011] UKPC 13, [2012] 2 AC 194.

136 Eg Gerard McCormack, 'The Liability of Trustees for Gross Negligence' [1998] Conveyancer and Property Lawyer 100; *Spread Trustee Co Ltd v Hutcheson* (n 135) [240]ff.

137 (n 125).

4 Application and utility of the two-stage approach

Finally, we aim to demonstrate the practical utility of the two-stage framework. One can look to taxonomy scholarship to identify relevant benchmarks. That scholarship suggests that certainty in the law and the reduction of costs is important,[138] as is ease of use, facilitating a critical overview to assist the analysis of the underlying field of law itself and providing normative guidance for decision-makers.[139] A critical overview has already been provided. As for the other matters, we show here how the two-stage framework makes the law easier to apply and errors harder to make and easier to spot.

The scenarios that come up most often are fiduciary joint ventures and directors' releases from liability. There are not many terribly relevant reported cases. However, this is not to be unexpected. *Hospital Products*, which solidified the idea that we can have fiduciary duties based on the peculiar facts of an engagement, was decided in 1984.[140] *Bristol and West Building Society v Mothew*, which, by holding that not every duty of a fiduciary was a fiduciary duty, bolstered the proposition that fiduciary duties can be scoped, was decided only in 1996.[141] Previously, fiduciary duties were thought in some quarters to be monolithic. Indeed, in its 1992 consultation paper, the Law Commission thought there were two approaches: 'status-based' and 'contract first'. In the former the contract could not modify or exclude the duties of a fiduciary where they were incompatible with the essence of the relationship, but in the latter contract terms would more readily relax the duty of loyalty.[142] The view that contractual terms can modify any instance of the duty of loyalty is new. However, the relevant factual scenarios are commonplace, and, while it will take time for decisions to emerge, emerge they should.

First, consider the error where judges have muddled the type of reasoning and applied stage one thinking in stage two or vice versa. The worst outcome is when this faulty reasoning leads to the wrong decision. The first instance judge in *Ross River* made this mistake.[143] He applied clause 10.5 in stage one as though it were a 'yes/no' sharp-edged contractual term – how it is applied at stage two – rather than as just one factor of many, concluding that it was incompatible with the existence of a fiduciary duty not to profit or to avoid conflicts, and thus some payments outside its specification were permitted.[144] His decision had to be reversed. The Court of Appeal correctly thought that the fiduciary rules were paramount. The two-stage framework assists in two ways. First, it indicates that clause 10.5's dominance was only at stage two, not stage one. Second, it makes it clear that fiduciary norms dominate and are not so easily displaced.

The next class of errors is where fiduciary norms have not been applied to stage two. Here, the judges applied contractual doctrines but not the additional fiduciary principles. It is possible that this is because they have implicitly accepted the idea that fiduciary duties are implied contractual terms, despite the post-*M&S v BNP* shift.[145] It is also possible that they have been led astray by the easy fit of the test for implication in stage two, which has led

138 Peter Birks, 'Equity in the Modern Law: An Exercise in Taxonomy' (1996) 26 University of Western Australia Law Review 1, 1.

139 Emily Sherwin, 'Legal Taxonomy' (2009) 15 Legal Theory 25, 39ff.

140 (n 5).

141 It appeared in the Law Reports somewhat later, in 1998: *Mothew* (n 132).

142 Law Commission, *Fiduciary Duties and Regulatory Rules* (Law Com CP 124, 1992) paras 3.3.9–3.3.10.

143 See above, text to n 21.

144 [2012] EWHC 81 (Ch) [259]; see generally [255]ff.

145 See (n 34).

them to think that it is the only test rather than just the first one.[146] In the following cases, the judges' reasoning fell into error but, more by luck than judgment, they reached the right decisions.

Global Container Lines Ltd v Bonyad Shipping Co (No 1) concerned a fiduciary shipping joint venture between the parties, who themselves operated shipping services independently.[147] There were two relevant geographic areas. The 'existing services area' comprised Global's existing Indian Ocean and Red Sea services, and the 'joint venture area' comprised services covering the rest of the Persian Gulf but excluding those in the existing services area. The basic issue was that Global, as a fiduciary, was prohibited from competing with its principal (the joint venture operation) unless clear consent was obtained. It was held that no competition was permitted in the joint venture area. This is not surprising – the duty of loyalty is inconsistent with the right to compete. Conversely, it was held that competition in the existing services area was impliedly consented to by Bonyad, and thus there the fiduciary duty was displaced altogether.

One may reach this route by applying our framework. As Rix J found, while express consent could not be construed from the various minutes in evidence before the court, it could be found through necessary implication. Bonyad had known about Global's extensive existing business and, in essence, giving up 51 per cent of its business and 75 per cent of its tramp business would have been 'unbusinesslike'.[148] The minutes Bonyad signed reflected the reality that the company knew and accepted that this business would continue, and this constituted consent.[149] One may go on to apply the regulating fiduciary rules. Bonyad knew the facts. The necessity was so strong, even resolving doubts against the profferer it was possible to infer subjective consent from it. The case went beyond *Kelly v Cooper*, which demonstrated only implied consent to act for two principals and demonstrated consent to compete in certain areas.

The problem is that, while Rix J started along this route, rather than considering the regulating fiduciary rules, he applied *Kelly v Cooper* rather mechanically. Competition was inherent to the business and as such 'duties . . . of . . . natural candidates for the status of fiduciaries . . . have to be tailored to the facts and circumstances'.[150] He thought that *Kelly v Cooper* was 'particularly instructive',[151] and the present case was 'if anything clearer cut' than it. The last point is certainly true, but the factual differences were not considered in full. The outcome of *Kelly v Cooper* was merely that a fiduciary was permitted to act for multiple principals. Moreover, estate agents act for multiple clients, but they do not positively compete with them. Self-sacrifice is inherent to a fiduciary relationship, and permitting the estate agent in *Kelly v Cooper* to act for multiple principals made considerably fewer inroads into this norm. If a case is to go further, one must explore why rather more fully.

Gamatronic (UK) Ltd v Hamilton concerned two directors' releases of liability for, inter alia, breaches of fiduciary duty.[152] This question was the scope of the release agreements.

146 Above, text near n 75.

147 (n 8).

148 Ibid 543.

149 Ibid 545.

150 *Global v Bonyad* (n 8) 545.

151 Ibid 545.

152 [2016] EWHC 2225 (QB). A curious question is why the Companies Act 2006, s 232, purportedly voiding such exclusions, was not litigated. As Ross Cranston, 'Limiting Directors' Liability: Ratification, Exemption and Indemnification' [1992] Journal of Business Law 197, 198 says, its relationship with ratification (authorisation) has not been clarified (considering its forerunners: Companies Act 1985, s 310; Companies Act 1948, s 205).

As a stage two issue, sufficiently informed consent is required, which is a subjective requirement,[153] but the principal was not aware of the relevant breaches of fiduciary duty when the releases were drawn up. The judge applied the contractual law of releases, for which the leading case is *Bank of Credit and Commerce International SA v Ali*.[154] He held that, since the principal was not aware of the relevant breaches when the release was made, it did not cover them.[155] This indeed is the right rule, but the problem is that it was reached without considering the fiduciary norms, and authority was applied without considering its applicability.

The complication is that the law of releases has become detached from its equitable roots. It once required the relevant matters to have been in the actual contemplation of the parties – that is, it was explicitly a subjective test.[156] Nowadays, contract doctrine holds that the test is objective. As a purely common law contract case (BCCI was not a fiduciary to Ali), the House of Lords rejected an inquiry into the parties' subjective states of mind: 'the court . . . makes an objective judgment based on the materials already identified'.[157] The Law Lords simply did not consider whether matters should be different for fiduciary liability. Perhaps they would have been sensitive to the question of fiduciary breach had it been put to them, and perhaps they would have held that equitable principles should have continued to apply to equitable duties. The judge in *Gamatronic* did not look behind *Ali*, and thus this issue was not considered. Our framework, specifically the sufficiency rule, makes the correct enquiry clear by expressly stating the subjectivity requirement.

The same error also occurred in an interim hearing in the same proceedings, where another first instance judge also thought *Ali* applied without considering the material differences between contract law and fiduciary law.[158] Again, in *John Youngs Insurance Services Ltd v Aviva Insurance Service UK Ltd*[159] and in *Nathan v Smilovitch*[160] the judges went straight to Ali without considering fiduciary matters. *John Youngs* was correctly decided because the release was held not to apply even on the less-stringent contract law principles; similarly, in *Nathan v Smilovitch*, the relevant term could not be implied even in contract.

It is surely inevitable that at some point a case will come up where the relevant term will be sufficient on contract law principles, but not on fiduciary law principles. Then, unless the fiduciary rules are applied, the wrong decision will be made.

5 Conclusion

The two-stage framework results inevitably when constructing doctrine from the fundamentals of fiduciary law, namely the purposes of protection and autonomy and that even the 'bad man' fiduciary needs some level of certainty. The need to uphold fiduciary standards and protect the principal demands a multi-factorial approach. The need for a

153 See above, text to (n 113).

154 (n 82).

155 *Gamatronic* (n 152) [156]ff.

156 *Lyall v Edwards* (1861) 6 Hurl & N 337, 158 ER 139; *Farewell v Coker* (1726) 2 Jac & W 192; 37 ER 599; (1726) 2 Mer 171, 354; 35 ER 905, 973.

157 *BCCI v Ali* (n 82) [8].

158 *Gamatronic (UK) Ltd v Hamilton (interim application)* [2013] EWHC 3287 (QB) [16]ff.

159 [2011] EWHC 1515 (TCC), [2012] 1 All ER (Comm) 1045. There is also *SPL Private Finance (PF1) IC Ltd v Arch Financial Products LLP* [2014] EWHC 4268 (Comm) [295] where Walker J held that 'This contention falls to be considered in accordance with well-established principles for determining the true meaning of contracts' in a fiduciary case.

160 [2002] EWHC 1629 (Ch) [13].

fiduciary to be fairly and safely remunerated demands a sharply logical ruleset. Since these requirements are fundamentally opposed, they can only exist in separate stages.

In coming to this conclusion, there has been one key thread running through the argument, namely the dominance of fiduciary principles. Thus, when contractual doctrines are compatible and useful, they can be and have been received into fiduciary law. However, one must not forget that contractual authorisation is predominantly fiduciary, and fiduciary principles take precedence over contractual principles. The duty of loyalty is not always moulded by the terms of the contract and does not always accommodate itself to them. This is easy to forget, and in some cases has been forgotten.

The courts appear to be inching towards this framework. That it has not yet been recognised explicitly is unsurprising in a system of case-by-case development in an immature part of the law, particularly given that the judge's primary task is to decide the instant case rather than construct theory. The post-*Hospital Products* era has been short, and what is going on appears to be an instance of what Llewellyn called 'slow-growing wisdom'.[161] As the common law – including equity – moves on, it creates a consistent and coherent body of law, even as it makes mistakes along the way.

161 K N Llewellyn, *The Bramble Bush: On Our Law and its Study* (Oceana 1951) 44.

NILQ 70(2): 263–280

Forfeiture of payment to a delinquent agent

PETER DEVONSHIRE*

Faculty of Law, University of Auckland

Abstract

A principal's claim for forfeiture of an agent's commission following the latter's default can be explained from both an equitable and common law perspective. In equity, the threshold for forfeiture is imprecise and reflects the vagaries of fiduciary obligations. At common law, forfeiture can be claimed in proceedings for money had and received. However, the principal's claim may be subject to counter-restitution in respect of the agent's services. It is argued that the common law provides an effective means of reconciling the parties' interests and that an approach based on failure of consideration would provide a more coherent and principled outcome in equity.

Keywords: agent; fiduciary; forfeiture of commission; money had and received; failure of consideration.

1 Introduction

An agent may receive a payment or benefit which is tainted by a breach of duty. This can take the form of a bribe or secret commission from a third party to induce an agent to betray his duty to his principal.[1] Such arrangements are manifestly corrupt and the agent is personally accountable to the principal for the proceeds.[2] A further consideration is that the agent may receive a commission or emolument from the principal for the performance of agreed services. If the agent dishonestly performs those services, he may be compelled

* Professor of Law, University of Auckland. I am indebted to Peter Watts and Rohan Havelock for their helpful comments on an earlier draft of this article.

1 Unless the context requires otherwise, the term 'agent' denotes an agent or employee and the term 'principal' denotes a principal or an employer.

2 *Fawcett v Whitehouse* (1829) 1 Russ & M 132 (Ch). See further *T Mahesan S/O Thambiah v Malaysia Government Officers' Co-Operative Housing Society Ltd* [1979] AC 374 (PC) 380. It is beyond the scope of this paper to discuss the implications of such accountability. It is generally the case that the agent's liability is both personal and proprietary. This is relatively uncontroversial in most Commonwealth jurisdictions. See, for example, *Grimaldi v Chameleon Mining NL (No 2)* [2012] FCAFC 6 at [510]; *Attorney General for Hong Kong v Reid* [1994] 1 AC 324 (PC). Historically, the position in England was less certain, but this has now been clarified by the UK Supreme Court in *FHR European Ventures LLC v Cedar Capital Partners LLC* [2014] UKSC 45, [2015] AC 250, which overruled the authority of *Lister & Co v Stubbs* (1890) Ch D 1 (CA) and affirmed that proprietary relief can be asserted in respect of the proceeds of bribery in the hands of a fiduciary.

to forfeit any remuneration to the principal. There are obvious distinctions between third-party bribes and commissions received directly from the principal. Notably, receipt of a bribe is a wrong per se whereas a commission paid by the principal is not.[3] Liability to forfeiture in the latter case arises from the wrongful performance of authorised acts.

This article will consider the elements of the principal's claim for forfeiture of remuneration following an agent's breach of duty. Other remedies can be invoked for breach of an agent's duties, such as rescission, damages, equitable compensation and account of profits. Reference will be made to these forms of relief, but a detailed treatment falls outside the scope of this article. Similarly, this article addresses conduct which constitutes a breach of fiduciary duty. A failure of duty in tort or a breach of contract (in respect of non-fiduciary performance) will not be considered.

The current state of the law is problematic. In this area, the conceptual bounds of equity and law are amorphous and, as a result, principles are often applied interchangeably. In addition, the Court of Appeal's 'spectacularly punitive reasoning'[4] in *Imageview Management Ltd v Jack*,[5] has attracted controversy and led to further uncertainty. This article will examine the status of an agent's commission where an agent engages in misconduct which strikes at the core of the principal–agent relationship. It will be asked whether the law in this area is being applied on a principled basis and whether there should be greater emphasis on the taxonomical status of the wrong. The latter will address the jurisdictional classification of an agent's default and the remedial implications which flow from it.

An agent's misconduct is explicable as both a breach of equitable duty and a common law wrong. The former is an obvious corollary of the agent's status as a fiduciary.[6] The latter is enforced by proceedings for breach of contract or an action for money had and received.[7] The intersection of law and equity raises intriguing issues as to the nature and taxonomy of the wrong and the remedial regimes that underpin it. From an equitable perspective, an element of dishonesty may be required for forfeiture of the principal's commission. This is not a universal rule and a lower threshold of equitable fraud[8] may suffice. From a common law perspective, the more egregious the wrong, the more readily the court will conclude that there is a total failure of consideration, rendering any payment liable to forfeiture. The jurisdictional basis of relief may in turn affect the form of the order. Discretionary relief in equity is potentially more nuanced and contrasts with the more absolute nature of forfeiture for money had and received.

3 Another obvious distinction is that by definition a secret commission is not disclosed to the principal.

4 P Watts, 'Forfeiture of Agents' Remuneration' in P Devonshire and R Havelock (eds), *The Impact of Equity and Restitution in Commerce* (Hart 2019) 203. For the facts of this case, see text at n 32.

5 [2009] EWCA Civ 63, [2009] Bus LR 1035. This judgment is insightfully reviewed in P Watts, 'Restitution and Conflicted Agents' (2009) 125 Law Quarterly Review 369.

6 *Crocs Europe BV v Anderson T/A Spectrum Agencies* [2012] EWCA Civ 1400, [24].

7 Restitution may be regarded as operating independently of the 'common law' wrong of breach of contract. From this viewpoint, money had and received is a discrete basis for recovery. See R Grantham and C Rickett, 'On the Subsidiarity of Unjust Enrichment' (2001) 117 Law Quarterly Review 273, where it is argued that the law of unjust enrichment serves to reverse certain transfers of wealth and that unjust enrichment fulfils a supplementary function, operating outside the law of contract and the law of wrongs.

8 'Equitable fraud' does not require actual dishonesty or deliberate wrongdoing. 'Fraud' in this context can be generally equated with unconscionable conduct. In *Nocton v Ashburton* [1914] AC 932 (HL), Viscount Haldane LC explained (at 954): 'In Chancery the term "fraud" thus came to be used to describe what fell short of deceit, but imported breach of a duty to which equity had attached its sanction . . . What it really means . . . is not moral fraud in the ordinary sense, but breach of the sort of obligation which is enforced by a Court that from the very beginning regarded itself as a Court of conscience.'

It is axiomatic that a remedy is a response to an obligation. Equity views an agent's breach primarily as the betrayal of a relationship of trust and secondarily as a failure to observe the duty which underlies that relationship.[9] Common law proceedings are founded on failure of consideration. The forfeiture of remuneration can be explained on either basis. As a result, the remedial response to an agent's breach vacillates between the normative paradigms of law and equity. This article will explore these issues and propound a modified approach based on common law doctrine.

At the outset, the central term 'forfeiture' should be explained. At issue is the status of an agent's payment, usually expressed as a commission. The term 'forfeiture' has two main applications. In the first, the principal seeks to recover a payment made to a defaulting agent. In the second, commission is unpaid and the principal may be entitled to resist the agent's claim for this sum.[10] In the latter setting, the term 'forfeiture' somewhat inaccurately describes a defence to the enforcement of a debt or other contractually agreed benefit. Both scenarios contemplate that the agent has provided some form of service.[11] If, however, no service has been rendered, the obligation to pay has simply not accrued. In the absence of performance there is no basis for an agent's claim to commission and by definition, no requirement to defend such a claim.

This article has five main parts. In 'Breach as an equitable wrong', forfeiture of remuneration is assessed as an equitable response to a breach of fiduciary duty. It is concluded that forfeiture does not fit readily within the remedial scheme of equity and cannot be satisfactorily explained as an account of profits or equitable compensation. The next part, 'Forfeiture and fiduciary doctrine', examines the variable nature of equitable duty and the absence of a unifying principle with respect to the role of forfeiture. This is followed by 'Common law perspective', where the taxonomy of the wrong is seen in broader focus. This introduces the influence of common law principles in defining the status of an agent's commission following default. It is argued that forfeiture should be confined to cases where there is a total failure of consideration. In this regard, the common law provides a coherent response to an agent's breach of duty and the circumstances in which remuneration can be forfeited. This is expanded in 'Equity and failure of consideration', where it is proposed that a common approach to the forfeiture of commission should be adopted in equity and law. A model based on failure of consideration would produce a more principled and consistent outcome in equity. The final substantive section, 'Jurisdictional and doctrinal implications of a unitary standard', examines some fundamental jurisdictional and doctrinal issues and concludes that neither unduly compromises the changes proposed in this article.

2 Breach as an equitable wrong

It is generally held that forfeiture of a commission flows naturally from breach of the agent's fiduciary duty.[12] This strict view draws from the exacting standards traditionally

9 This is aptly captured in Lord Russell's pronouncements on the strictness of the non-profit rule. See *Regal (Hastings) Ltd v Gulliver* [1967] 2 AC 134n (HL), 144–45.

10 For example, in the leading case of *Imageview* (n 5), an agent in breach of fiduciary duty to his principal lost the right to future agency fees in addition to remuneration that had already been paid.

11 The principles discussed in this paper apply in either situation. See *Andrews v Ramsay & Co* [1903] 2 KB 635 (DC) 638; *Imageview* (n 5) [51].

12 This conclusion is readily drawn where there is an absence of good faith and the breach is not trivial. See *Keppel v Wheeler* [1927] 1 KB 577 (CA), 592; *Andrews v Ramsay* (n 11) 636–38; *Imageview* (n 5) [47]–[50]; *Stevens v Premium Real Estate Ltd* [2009] NZSC 15, [2009] 2 NZLR 384, [89] and [90].

imposed on fiduciaries. The basic obligations of a fiduciary are enshrined in the profit and conflict rules. By the former, the fiduciary is not permitted to profit from her office, save with the informed consent of the principal.[13] By the latter, the fiduciary must not place herself in a position where personal interest conflicts with duty to the trust.[14] The profit and conflict rules operate regardless of whether the defaulting fiduciary acted in bad faith or consciously committed any wrongdoing. The mere fact of being placed in a position where duties are, or may be compromised, is sufficient.[15] The fiduciary is accountable whether or not the principal has suffered any loss.[16] Nor does it matter that the principal would not or could not obtain the profit.[17]

These core principles define equity's expectations of fiduciary duty. It is generally held that results – in this case remedial outcomes – are fashioned by principles and normative objectives. Equity's prescriptions are empowered by a close relationship between principle and remedy. In this regard it is uncontroversial that equitable remedies are 'more elastic'[18] than remedies at law and more adequately reflect the policies they enforce. Speaking of equitable relief for fiduciary wrongs, it has aptly been said that:[19]

> . . . [equitable] remedies will be fashioned according to the exigencies of the particular case so as to do what is 'practically just' as between the parties.

The unexplored question is the extent to which remedy enforces doctrine and the juristic basis on which this is achieved. There are several permutations. The forfeiture of remuneration following an agent's breach of fiduciary duty[20] can be explained as an account of profits, equitable compensation,[21] or restorative relief. Each will be considered in turn.

An account of profits is perhaps the least satisfactory rationale for forfeiture of commission.[22] In claiming an agent's commission, the principal is recovering his own funds as a result of the agent's repudiatory conduct. A commission received by a false fiduciary should be categorised as a wrongful receipt rather than a profit. Semantically, at least, this would suggest that the remedy has more affinity with an action for money had and received than disgorgement. The distinction is apparent through the modern law of unjust enrichment, which would characterise the gain as an unjust enrichment by subtraction.[23] The remedial response is reversal of a transfer of wealth from the principal to the agent.[24]

13 *Boardman v Phipps* [1967] 2 AC 46 (HL) 105; *Warman International Ltd v Dwyer* (1995) 182 CLR 544, 554; *Gwembe Valley Development Co Ltd v Koshy* [2003] EWCA Civ 1048, [2004] BCLC 131, [43]–[45].

14 *Bray v Ford* [1896] AC 44 (HL) 51.

15 *Regal (Hastings) Ltd* (n 9); *Chan v Zacharia* (1984) 154 CLR 178.

16 *Boardman v Phipps* (n 13).

17 *Howard v Commissioner of Taxation* [2014] HCA 21, (2014) 253 CLR 83.

18 *Nocton v Ashburton* (n 8) 952 per Viscount Haldane LC.

19 *Maguire v Makaronis* [1997] HCA 23, (1997) 188 CLR 449, 496 per Kirby J.

20 As noted above, similar reasoning applies in respect of restoring payments made by the principal and resisting the agent's claim for payment where no payment has been made. See, for example, *Rhodes v Macalister* (1923) 29 Com Cas 19 (CA).

21 *Hurstanger Ltd v Wilson* [2007] EWCA Civ 299, [2007] 1 WLR 2351, [48].

22 An account can also be invoked where the principal seeks to rescind an agent's contract. Here, an account of profits is instrumental in adjusting the parties' rights to achieve *restitutio in integrum*. In addition, an account may operate more expansively to replicate the effect of rescission where restitution *in specie* is not available. See *Mahoney v Purnell* [1996] 3 All ER 61 (QB), 88–89; *Halpern v Halpern* [2007] EWCA Civ 291, [2008] QB 195, [60].

23 P Birks, *An Introduction to the Law of Restitution* (Clarendon 1989) ch 1.

24 A Burrows, *Remedies for Torts and Breach of Contract* (2nd edn, Oxford University Press 1994) 287.

Equitable compensation does not strictly reflect the elements of forfeiture. By definition, the purpose of equitable compensation is to compensate for loss. This is problematic in the present context because equitable compensation or damages will usually place the injured party in the position he would have been in if the agent's services had been faithfully performed.[25] If loss has been fully compensated, the claim to any additional sum must be understood on a different basis. Forfeiture of commission has been explained as restorative relief. As such it is distinct from an account of profits, equitable compensation and damages. Thus, it has been held that compensatory and restorative awards are cumulative and the principal is not required to exercise an election.[26] Both remedies may therefore be claimed in respect of the same wrong.[27]

This is aptly demonstrated in *Stevens v Premium Real Estate Ltd*,[28] where the vendors were induced to sell their property for a reduced sum as a result of false representations by the purchaser. The vendors' estate agent was complicit in the deception and actively misled his clients, the vendors. Not surprisingly it was held that the estate agent was in breach of fiduciary duty. The estate agent was ordered to refund the commission as well as compensating the vendors for their loss. The Supreme Court of New Zealand elaborated:[29]

> They [the plaintiffs] paid the commission on the premise that Premium [the defendant] had faithfully performed its contractual and fiduciary duties. Premium had not done so. Whether the commission should be refunded as restorative damages in addition to compensatory damages does not . . . involve any question of election . . . The policy which lies behind compensatory damages is simple. It is to compensate for loss caused by civil wrongs. The policy behind restorative damages [includes] . . . reinforcing fiduciary obligations by removing from the defaulting fiduciary a sum paid pursuant to a contract which has not been faithfully performed.

If equitable compensation and restorative awards are cumulative, it follows that, unless the latter is duplicative, it does not fulfil a compensatory function. It is somewhat strained to characterise a commission as a loss to the principal. The agent's breach of duty does not necessarily denote that the principal's contract was not performed or that losses flowing from the breach cannot be fully recovered through equitable compensation.

For example, in *Stevens v Premium Real Estate Ltd*, the Supreme Court awarded damages representing the difference between the actual sale price and the market value of the plaintiffs' property. In addition, the agent's commission was forfeited because it had not been earned by good faith performance of its services.[30] In the result, the plaintiffs were made whole by the award of compensation and did not sustain any further financial loss. In monetary terms the consequences of the wrong were expunged. The additional forfeiture of commission was not compensatory and must be rationalised on a different basis.

25 See discussion in this section, below, and section 4 'Common law perspective'.

26 However, it is clearly established that an account of profits and equitable compensation are alternative remedies and the claimant will be put to an election. See *Tang Man Sit v Capacious Investments Ltd* [1996] 1 AC 514 (PC).

27 G Virgo, *Principles of the Law of Restitution* (2nd edn, Oxford University Press 2006) 461–63.

28 *Stevens* (n 12).

29 Ibid [107] per Tipping J. See also [12], [30]–[31] per Elias CJ and [89]–[90] per Blanchard, McGrath and Gault JJ. See further *Andrews v Ramsay* (n 11) 637–38; *Imageview* (n 5) [49] –[51].

30 The Supreme Court considered that the agent's conduct had fundamentally undermined the parties' relationship (*Stevens* (n 12) [31], [89], [90], [94] and [109]).

It can be argued that restorative awards reflect the view that an agent's remuneration is a contractual benefit for a service that has not been performed. Put compendiously, a service performed corruptly is no service at all. In a narrow and somewhat artificial sense, in paying for that service, the principal has sustained a loss. Despite its theoretical allure, this is unconvincing. For example, in *Stevens v Premium Real Estate* the plaintiffs were restored to the economic position they would have enjoyed if the wrong had not occurred. Adjusted by damages, the plaintiffs' property had been sold at market value.[31] In fact the forfeiture of commission placed the plaintiffs in a superior position as a result of the breach because the defendant's services in effecting the sale were unremunerated.

A similar result can be seen in the influential decision of *Imageview Management Ltd v Jack*.[32] The plaintiff was retained by the defendant, a professional footballer, to act as his agent for an agreed commission, to obtain employment with a UK football club. The plaintiff negotiated a contract for the defendant to play for a Scottish football club. As the defendant was a national of Trinidad and Tobago, he required a work permit. The football club offered to pay the plaintiff £3000 for obtaining a work permit. The plaintiff did not disclose this agreement to the defendant. Subsequently the defendant discovered these arrangements and stopped paying the commission under the agency agreement. The plaintiff commenced proceedings to recover the unpaid commission and the defendant counterclaimed for the return of commission that had already been paid as well as the £3000 fee the plaintiff had received from the club.

It was held at first instance and on appeal that the plaintiff had a conflict of interest with respect to the payment from the club and had acted in breach of fiduciary duty. The defendant was therefore relieved of liability with respect to future commissions and was entitled to recover commission already paid as well as the defendant's £3000 fee. As in *Stevens v Premium Real Estate*, the agent forfeited any contractual entitlement to payment under the agency agreement. The forfeiture of commission in *Imageview* meant that the defendant had benefited from the agent's services in obtaining a work permit and the latter was unpaid for this service.

In sum, restorative awards do not necessarily serve to redress the consequences of an unperformed contract.[33] Moreover, there are conceptual difficulties in classifying a restorative order as either gain-based or compensatory. At the same time, despite concerns as to the provenance of the wrong and the nature of the order, it seems counterintuitive to disclaim equitable jurisdiction in respect of forfeiture.[34] As noted, in *Stevens* and *Imageview*, the outcome was advantageous to the plaintiff who was fully compensated for loss and also relieved of financial obligation in respect of the defendant's services. This is not inconsistent with the broad objectives of equity. It is trite that the deterrent principle animates equity's governance of fiduciary relationships. As a simple dogma, the forfeiture of commission gives expression to this principle by ensuring that an errant fiduciary is stripped of the benefits of that relationship and the temptation to abuse the trust reposed in him.

31 In addition, the plaintiffs were treated favourably in that there was evidence that in negotiating the terms of sale the plaintiffs were prepared to accept less than market value.

32 *Imageview* (n 5). This judgment is insightfully reviewed in Watts, 'Restitution and Conflicted Agents' (n 5).

33 See below, section 4 'Common law perspective', which proposes that remuneration can be restored to the principal if there is a total failure of consideration.

34 *Stevens v Premium Real Estate* and *Imageview* were determined by reference to an agent's breach of fiduciary duty. In both cases the remedies were a response to an equitable wrong. See *Imageview* (n 5) [50]; *Stevens* (n 12) [30] (Elias CJ), [90] (Blanchard, McGrath and Gault JJ) and [109] (Tipping J).

The scope of the deterrent principle is not without controversy but its application remains true in the present context. While some academic commentators question the status of deterrence, the forfeiture of commissions is usually not directly in issue, or at best, on the margin of such debate. Notably, Lionel Smith argues that the profit rule does not have any prophylactic or deterrent function and that a fiduciary's accountability for illicit gains is more appropriately explained as a primary rule of attribution.[35] While the fiduciary's misconduct can invariably be classed as a wrong, accountability to the principal for such gains is merely a vindication of the latter's primary right to such receipts.[36] Effectively, the errant fiduciary lacks capacity to acquire beneficial receipt. Such reasoning is applicable to unauthorised gains from third-party sources, but less obviously so in respect of agreed commission payments from the beneficiary. Such payments are not an unauthorised receipt in the hands of the fiduciary. Nor can it be said that the fiduciary lacked capacity to obtain such receipts.

It may be countered that deterrence does not capture the true dynamics and that forfeiture is more obviously categorised as punitive. Orthodox reasoning suggests that equity embraces the former and eschews the latter. While equity has historically struck down penalties,[37] the objective of deterrence has long informed equity's remedial regime. Notwithstanding the traditional disavowal of penal jurisdiction,[38] it may be questioned whether such reasoning is more semantic than substantive. Although conceptual distinctions can be drawn between punishment and deterrence, the distinctions are less apparent in their application.[39] At a basic level, sanctions reflecting deterrence will have prejudicial consequences to the wrongdoer. It is but a slight step to suggest that in practical terms at least, some deterrent-inspired remedies have a punitive effect.[40] However, a more precise distinction can be drawn in relation to the forfeiture of payment to a delinquent agent. It is submitted that forfeiture is clearly penal if the defendant is stripped of a commission when the principal has received the benefit of the agent's services and been fully compensated for any loss.[41] In such cases, equitable relief appears to exceed the bounds of compensation or disgorgement of gains, and it is preferable to openly recognise that the objective is punitive. Conversely, forfeiture is not so obviously penal if the agent has failed to perform the agreed service or acted so corruptly that the principal has not derived any benefit from those services. In this setting, restoring to the principal the value of an unearned commission is little more than a rebalancing of the economic realities of the transaction.

35 Lionel Smith, 'Deterrence, Prophylaxis and Punishment in Fiduciary Obligations' (2013) 7 Journal of Equity 87; Lionel Smith, 'Fiduciary Relationships: Ensuring the Loyal Exercise of Judgement on behalf of Another' (2014) 130 Law Quarterly Review 608.

36 Smith, 'Deterrence, Prophylaxis and Punishment' (n 35) 100; Smith, 'Fiduciary Relationships' (n 35) 628–29.

37 In Somers J's concise utterance: '[E]quity and penalty are strangers' (*Aquaculture Corp v New Zealand Green Mussel Co Ltd* [1990] 3 NZLR 299 (CA), 302). See also *Harris v Digital Pulse Pty Ltd* [2003] NSWCA 10, (2003) 56 NSWLR 298, [338] per Heydon JA.

38 *Vyse v Foster* (1872) LR 8 Ch App 309, 333. See also *Dart Industries Inc v Décor Corporation Pty Ltd* (1993) 179 CLR 101, 111 and 114.

39 Indeed, the terms 'deterrence' and 'punishment' are often used interchangeably. See, for example, *Gray v Motor Accident Commission* (1998) 196 CLR 1, [15].

40 Whether this is characterised as a slight step or a major step epitomises the doctrinal rift between Mason P and Heydon JA in *Harris v Digital Pulse* (n 37). See further P Devonshire, 'Account of Profits for Breach of Fiduciary Duty' (2010) 32 Sydney Law Review 389.

41 See discussion above regarding *Stevens* (n 12).

The next section will examine the variable nature of equitable duty, the absence of unifying principle with respect to the role of forfeiture and the uncertain status of an agent's commission following default.

3 Forfeiture and fiduciary doctrine

If forfeiture claims are governed by fiduciary doctrine, then, except in trivial cases,[42] it might be expected that a conflict of interest or breach of the profit rule will be sufficient for an agent to forfeit all remuneration. In *Keppel v Wheeler*[43] the Court of Appeal was uncompromising on the point:[44]

> Now I am quite clear that if an agent in the course of his employment has been proved to be guilty of some breach of fiduciary duty, in practically every case he would forfeit any right to remuneration at all.

The vigour of this principle was undiminished in the later case of *Imageview Management Ltd v Jack*,[45] where Jacob LJ proclaimed:[46]

> The law imposes on agents high standards . . . An agent's own personal interests come entirely second to the interest of his client. If you undertake to act for a man you must act 100%, body and soul, for him. You must act as if you were him. You must not allow your interest to get in the way without telling him.

Following a review of the leading authorities, his Lordship concluded that: 'Once a conflict of interest is shown . . . the right to remuneration goes.'[47] Such views suggest that forfeiture is a natural consequence of transgressing fiduciary duty and that once the duty is breached, the remedy applies.[48] By invoking the strict standards by which fiduciaries are governed,[49] it can be inferred that a low threshold of equitable fraud will suffice for forfeiture of a commission. However, this is deceptive. Equitable duty is not absolute.[50] The fiduciary principle embraces many forms of relationship and the nature of fiduciary duty will be driven by the particular facts. For example, whilst estate agents owe fiduciary duties to their principal, the obligations are fashioned by the terms of their engagement and the commercial context in which the parties operate. Thus, it is recognised that conflict between duty and duty is an unavoidable and permissible aspect of an estate agent's business.[51] Estate agents can act for a number of principals, some of whom may be competing with each other in the same market. By definition, the interests of these parties conflict and a common agent will in turn be conflicted in his duty to each.[52]

Moreover, forfeiture of a commission is but one of a number of remedial responses to an agent's breach of duty. Despite emphasising the strict nature of equitable

42 *Imageview* (n 5) [44].

43 *Keppel v Wheeler* (n 12).

44 Ibid 592 per Atkin LJ. But cf text below, where on the facts it was held that the agent was entitled to retain an earned commission.

45 *Imageview* (n 5).

46 Ibid [6] per Jacob LJ.

47 Ibid [44].

48 *Stupples v Stupples & Co (High Wycombe) Ltd* [2012] EWHC 1226 (Ch), [25].

49 *Boardman v Phipps* (n 13); *Regal (Hastings) Ltd* (n 9).

50 As the High Court of Australia stated in *Howard v Commissioner of Taxation* (n 17) (at [34] per French CJ and Keane J): 'Despite their broad judicial formulations fiduciary duties are not infinitely extensible.'

51 *Kelly v Cooper* [1993] AC 205 (PC). A corollary is that an agent is not required to share confidential information between principals whose interests may conflict (214–15).

52 In most settings this will constitute a breach of fiduciary duty. See further Millett LJ in *Bristol & West Building Society v Mothew* [1998] Ch 1 (CA), 18.

obligations, courts have equivocated as to the extent of relief where the breach is only incidental to the fiduciary's duties or, in Jacob LJ's memorable epithet, an instance of 'harmless collaterality'.[53] It is here that the taxonomy of the wrong is seen in broader focus. Thus far, forfeiture has been characterised as a response to an equitable wrong. This is a logical point of departure given that the relationship of principal and agent is fiduciary. However, agency relationships are usually contractual and accordingly common law principles may have a bearing on liability and remedies for breach.[54] This is particularly evident in the commercial sphere where the more sophisticated the transaction, the less compelling the case for equitable intervention.

From this perspective, breach of a duty of loyalty does not automatically confer a right of forfeiture.[55] Loyalty is a cornerstone of fiduciary duty[56] and – from an equitable standpoint at least – it seems unduly diffident to deny a full range of sanctions for breach of a fundamental duty. However, viewed in its contractual matrix, disloyalty may not necessarily go to the heart of the relationship. While this may be untenable with respect to an express trustee,[57] it has considerable force in other settings, where fiduciary duty is malleable in content and emphasis.[58] Here, the breach and the remedial consequences are a question of degree. In some circumstances liability may be avoided altogether or misconduct may be sufficiently redressed by a compensatory award. Moreover, if the court orders rescission on terms, the status of an agent's remuneration may be less bound up with unitary factors such as failure of consideration and more directed to a general adjustment of the parties' interests.

In determining a claim for forfeiture of an agent's commission, the question is not whether an agent has been disloyal, but whether an agent has been disloyal fundamentally. A key factor is whether the agent has acted dishonestly.[59] If the breach of duty falls short of dishonesty, the agent may be allowed to retain payment for services. In *Keppel v Wheeler*,[60] the plaintiff engaged the defendant estate agents to sell a property. The property was listed at £6500, but the plaintiff indicated that he would accept £6000. The defendants introduced a purchaser who offered £6150. The plaintiff accepted this offer subject to contract. Before the contracts were exchanged, the defendants received a higher offer of £6750. At that point the plaintiff was in a position to reject the first offer and accept the second. However, instead of communicating the second offer to the plaintiff, the defendants approached the original purchaser with the suggestion that he acquire the subject property and onsell it for a profit to the second offeror. The

53 *Imageview* (n 5) [44].

54 It is widely accepted that equitable relief may be marginalised or excluded entirely by the terms of the parties' agreement: *Hospital Products Ltd v United States Surgical Corp* (1984) 156 CLR 41, 97. See also *John Alexander's Clubs Pty Ltd v White City Tennis Club Ltd* [2010] HCA 19, (2010) 241 CLR 1, [91]; *Henderson v Merrett Syndicates Ltd* [1995] 2 AC 145 (HL), 206.

55 *Wright Hassall LLP v Horton Jr* [2015] EWHC 3716 (QB), [59].

56 *Bristol & West v Mothew* (n 52) 18; *Beach Petroleum NL v Kennedy* [1999] NSWCA 408, (1999) 48 NSWLR 1, [196]–[202]; *Arklow Investments Ltd v Maclean* [2000] 1 WLR 594 (PC), 598–600; *Chirnside v Fay* [2006] NZSC 68, [2007] 1 NZLR 433, [15] and [78].

57 It has been forcefully argued that in respect of trustees there is an irreducible core of fiduciary obligations which cannot be abrogated (*Armitage v Nurse* [1998] Ch 241 (CA), 253–54).

58 The point can be pressed further. Not every relationship of trust and dependence is necessarily fiduciary. The status of commercial joint ventures, for example, hovers between law and equity. See *Gibson Motorsport Merchandise Pty Ltd v Forbes* [2006] FCAFC 44, (2006) 149 FCR 569, [13], and *Arklow Investments v Maclean* (n 56).

59 This contrasts with the more rigorous treatment of an 'innocent' breach by a trustee or a fiduciary in service of a trust. See *Boardman v Phipps* (n 13).

60 *Keppel v Wheeler* (n 12).

defendants acted in the mistaken belief that their duty to the plaintiff had been discharged when the purchaser accepted the first offer, subject to contract. In fact the defendants' duty subsisted until contracts had been formally exchanged.

The Court of Appeal held that the defendants were in breach of duty and awarded damages for the difference between the two offers. However, the defendants' counterclaim for commission on the sale was upheld. In contrast to proceedings for breach of trust, the court took account of the nature of the wrong and the defendants' subjective understanding of their obligations to the plaintiff. Whilst acknowledging the general principle that an agent is liable to forfeit remuneration for a breach of fiduciary duty, Atkin LJ conceded that:[61]

> [T]here may well be breaches of duty which do not go to the whole contract, and which would not prevent the agent from recovering his remuneration.

In finding for the defendants, his Lordship was mindful that the agents had acted in good faith and that the principal had received the benefit of the completed transaction.[62] A similar approach is evident in *Robinson Scammell & Co v Ansell*,[63] where an agent disclosed information regarding his client (the plaintiff) to a potential purchaser. Although this did not cause the purchaser to withdraw from the transaction, the agent was in breach of duty to the principal. Nevertheless, the agent was entitled to retain its commission as it had not acted dishonestly or in bad faith.[64]

What emerges from this line of authority is that the finding of a breach of duty does not conclusively determine the status of a commission. The courts press further. Attention turns to performance, focusing on the nature of the wrong and the agent's motives in pursuing a particular course of action. Forfeiture is triggered by an agent's failure to meet certain proscriptive duties,[65] which are principally expressed in terms of prohibition against receiving secret commissions[66] and procuring a sale to oneself or a company in which one is interested.[67] Again, fraud in its different manifestations, such as deceit, bribery and misappropriation, fall in this category.

4 Common law perspective

The previous section has identified different equitable perspectives to forfeiture of commission. Not uncommonly, however, concepts are applied interchangeably without regard to their jurisdictional origins. Important implications flow from the distinct focuses of law and equity. Attention now turns to the common law.

As noted, the primary equitable remedies for breach of fiduciary duty are an account of profits, equitable compensation and rescission. Where an agent betrays the interests of

61 Ibid 592 per Atkin LJ.

62 Ibid. In similar vein, Bankes LJ commented (at 588): 'an agent might quite properly claim his commission, and yet have to pay damages for committing a bona fide mistake which amounts to a breach of duty'.

63 *Robinson Scammell* [1985] 2 EGLR 41 (CA).

64 This approach was approved, obiter, in *Kelly v Cooper* (n 51), where the Privy Council opined that, even if a breach of fiduciary duty is established, the defendants would not lose their right to commission unless they had acted dishonestly.

65 From a contractual perspective, an agent's breach may relieve the principal of an obligation to pay commission. Here, the language of forfeiture is less appropriate and the dynamics are better expressed in terms of the non-accrual of the principal's obligation to pay an agreed sum to the agent.

66 *Stupples v Stupples* (n 48) [21].

67 *Robinson Scammell* (n 63).

his principal, he is personally accountable for any gain.[68] If the breach occasions loss to the principal, this may be recovered as equitable compensation. As cases like *Stevens v Premium Real Estate* demonstrate: (i) the contract may have been performed notwithstanding the agent's breach of duty, and (ii) any losses flowing from the breach may be fully recovered through equitable compensation. If the principal has not suffered any further economic harm, the additional sanction of forfeiture is unwarranted. It is submitted that forfeiture should be limited to cases where there is a total failure of consideration. In this regard, equity should follow the law by vindicating the principal's performance interest in the agent's services.

This must be placed in context. Forfeiture of an agent's remuneration can be explained on a different basis from disgorgement and compensation.[69] Failure of consideration involves a factual determination, which in one respect at least, stands outside the normative expectations of fiduciary duty. The general threshold for fiduciary liability is satisfied if there is a perceived risk of disloyalty.[70] It is sufficient if there is a 'significant'[71] or 'real sensible possibility'[72] of conflict. The strictness of this rule is reinforced by the court's reluctance to engage in speculation as to what might or might not have happened in different circumstances or in the absence of a breach of duty.[73] With respect to forfeiture, common law is only concerned with the fact, not the risk, of a breach of duty. A total failure of consideration denotes the required quality of the breach, hence insistence on a fundamental failure that effectively negates performance. On the approach argued in this article, equity's traditional sanctions in respect of the profit and conflict rules[74] are unaffected, but forfeiture should be aligned with the common law's insistence on a failure of consideration.[75]

This is thrown into focus when an agent's misconduct is assessed from a contractual perspective. As Matthew Conaglen notes, where remuneration for contracted services is received directly from the principal, *prima facie* these are authorised receipts in the hands of the fiduciary. This raises performance issues directed to whether the fiduciary's breach goes to the root of the contract of retainer.[76] Other commentators have noted the misalignment of law and equity in this setting. Seb Oram observes that forfeiture is often consistent with contractual analysis, particularly where there is no performance by the agent.[77]

68 For example, in the case of bribes: *Lister v Stubbs* (n 2); *Attorney General for Hong Kong* (n 2); *FHR European Ventures v Cedar Capital* (n 2).

69 See section 2 'Breach as an equitable wrong', above.

70 *Aberdeen Railway Co v Blaikie Brothers* (1854) 1 Macq 461 (HL), 471; *Boardman v Phipps* (n 13) 103; *Warman* (n 13) 557.

71 *Chan v Zacharia* (n 15) 199.

72 *Boardman v Phipps* (n 13) 124; *Queensland Mines Ltd v Hudson* (1978) 18 ALR 1 (PC), 3–4.

73 *Brickenden v London Loan & Savings Co* [1934] 3 DLR 465 (PC), 469; *Gwembe Valley Development* (n 13) [145]; *Murad v Al-Saraj* [2005] EWCA Civ 959, [2005] WTLR 1573, [105]ff.

74 The latter includes conflict between duty and duty.

75 There is an intermediate point between the common law approach to forfeiture and the equitable standard. If a fiduciary's services are performed negligently or improperly, forfeiture of remuneration is not in issue. Such disputes are governed by analogy to tort (*Bristol & West Building Society v Mothew* (n 52)). The proceedings are premised on performance, albeit defective performance, in respect of which compensation may be awarded.

76 M Conaglen, 'Remedial Ramifications of Conflicts between a Fiduciary's Duties' (2010) 126 Law Quarterly Review 72, 77 and 78.

77 S Oram, 'Forfeiture of Fiduciary Remuneration Following Breach of Duty: From Contract to Conscience' [2010] Lloyd's Maritime and Commercial Law Quarterly 95.

Developing this point, Peter Watts explains that forfeiture of remuneration can be understood from the common law perspective of money had and received, where the agent's misconduct constitutes a total failure of consideration.[78] However, the concept of 'total' failure is problematic, particularly if, as in *Stevens v Premium Real Estate*, the agent's services have been performed and damages have been assessed on that basis. Indeed, it is arguable that damages and restitution are inconsistent remedies because restitution of contractual consideration removes the basis on which the plaintiff is entitled to require the defendant to perform her contractual obligations.[79] In any event, leaving aside the status of damages, the prevailing view (despite criticism)[80] is that a failure of consideration must be total, not partial.[81] An agent's commission is usually an element of an entire contract. It is therefore an all or nothing situation. Partial failure will not suffice unless the contract is severable.[82] This is unlikely in a setting where corrupt performance cannot be segregated from the overall transaction.

It is therefore necessary to confront the fundamental question whether a breach of duty amounts to a total failure of consideration. From an equitable perspective, corrupt service may be so anathema to the concept of fiduciary duty as to effectively nullify any recognition of beneficial service. This is essentially a total failure of consideration in a different guise. However, equity's position is unpredictable. As previously noted, the fiduciary principle embraces many forms of relationship and the nature of fiduciary duty and its breach is driven by the particular facts. In some cases the threshold for forfeiture is equitable fraud; in other cases, forfeiture is only available if dishonesty and actual fraud are established. In contrast, the common law insists on a breach of basic proscriptive duties. It is evident that this is a more coherent and principled approach when considered from the perspective of a failure of basis.

Discussion to this point has adopted the traditional language of failure of consideration. However, in the present context it has aptly been observed that the term 'failure of basis' may be preferable to 'failure of consideration' because it more accurately depicts the essence of a claim directed to failed performance[83] and comprehends 'the state of affairs which was within the contemplation of the parties as the basis of their dealings'.[84] In contractual cases involving an agent's commission, this is directed to the nature of the agent's promised counter-performance.[85] A claim of failure of basis applies to fundamental breaches of the relationship. This would rule out narrow and technical infractions and reflect the common law's enforcement of basic prescriptive duties.[86]

78 Watts, 'Restitution and Conflicted Agents' (n 5) 369–71.

79 *Baltic Shipping Co v Dillon* (1993) 176 CLR 344, 355.

80 See discussion in G Virgo, *The Principles of the Law of Restitution* (3rd edn, Oxford University Press 2015) at 325–329.

81 *Fibrosa Spolka Akcyjna v Fairbairn Lawson Combe Barbour Ltd* [1943] AC 32 (HL), 77; *Baltic Shipping v Dillon* (n 79) 355–56; *Raulfs v Fishy Bite Pty Ltd* [2012] NSWCA 135, [88] and [89]. See further J Edelman and E Bant, *Unjust Enrichment* (2nd edn, Hart 2016) 266–69.

82 In this limited situation, restitution is available in respect of a discrete aspect of failed performance. See *Roxborough v Rothmans of Pall Mall Australia Ltd* [2001] HCA 68, (2001) 208 CLR 516, [19], [106] and [165]–[167]; *Barnes v The Eastenders Cash & Carry plc* [2014] UKSC 26, [2015] 1 AC 1, [114].

83 *Barnes v The Eastenders* (n 82).

84 *Roxborough v Rothmans* (n 82) [17] per Gleeson CJ, Gaudron and Hayne JJ. See similarly *David Securities Pty Ltd v Commonwealth Bank of Australia* (1992) 175 CLR 353, 382.

85 Failure of basis is not limited to failure of contractual performance and has wider application to a fundamental failure of a state of affairs (*Roxborough v Rothmans* (n 82) [16] and [104]).

86 See text following n 62 above.

If the principal seeks to recover an agent's commission, there is a corresponding obligation to effect counter-restitution.[87] As pithily expressed in *Ovidio Carrideo Nominees Pty Ltd v The Dog Depot Pty Ltd*:[88] 'an entitlement to sue for "counter restitution" is pro tanto an answer to a claim for restitution'. At first blush, it seems paradoxical and (from the principal's perspective) self-defeating for a delinquent agent to offset forfeiture with a claim for the value of corrupt services.[89] The most logical – albeit circuitous – response is that no benefit was conferred by those services due to a failure of basis.[90] The efficacy of this position will depend on the substantive effect of the breach. It is a question of degree. The principal's obligation to effect counter-restitution is nullified if an agent failed to confer any form of benefit. For example, if the promised service was not performed or was performed so corruptly as to negate any benefit to the principal. Thus, the scope of counter-restitution must be tested against the defence of good consideration.

It has been noted that restitution of benefits transferred from the principal to an agent may be conditional on an award to the agent for the fair value of services rendered. As explained by Edelman and Bant: 'The rationale of the defence [of good consideration] is that the plaintiff must not be unjustly enriched by an order for restitution.'[91] Thus, a defence of good consideration may limit or extinguish primary liability to effect restitution due to a failure of consideration.[92]

A defence of good consideration reverses[93] and complements unjust enrichment through restitution by addressing the substantive effect of the payee's services. This critical enquiry determines the retention and forfeiture of an agent's remuneration. Returning to the facts of *Stevens v Premium Real Estate*, the defendant was required to restore commission received from the plaintiffs notwithstanding that the plaintiffs were fully compensated for loss attributable to the purchaser's deception and placed in the same position as if the agent's services had been duly performed. The case for counter-restitution was strong. The defence of good consideration was evident on the facts and affirmed by the form of relief granted. In the next section it is argued that this approach should inform equitable doctrine with respect to forfeiture.

5 Equity and failure of consideration

It is submitted that an approach based on failure of consideration would produce a more consistent and predictable outcome in equity. In this setting, equitable relief should be

87 Counter-restitution is not confined to contractual relationships. See, for example, *Pavey & Matthews Pty Ltd v Paul* (1986) 162 CLR 221; *Ovidio Carrideo Nominees Pty Ltd v The Dog Depot Pty Ltd* [2006] VSCA 6, (2006) V ConvR 54–713.

88 *Ovidio Carrideo v The Dog Depot* (n 87) [49] per Nettle JA.

89 Correspondingly, allowances are granted in the court's equitable jurisdiction. In *Stupples v Stupples* (n 48) doubts were expressed, *obiter*, that a defaulting agent could receive an allowance in lieu of contractual commission (at [26]). See also *Avrahami v Biran* [2013] EWHC 1776 (Ch), [345].

90 The counter-argument of good consideration is discussed below.

91 Edelman and Bant (n 81) 365. The case of *Ovidio Carrideo v The Dog Depot* (n 87) 54–713 is instructive on this point. See further *Adrenaline Pty Ltd v Bathurst Regional Council* [2015] NSWCA 123.

92 *ANZ Banking Group Ltd v Westpac Banking Corporation* (1988) 164 CLR 662, 673; *Plan B Trustees Ltd v Parker* [2013] WASC 216, [87]–[91].

93 As good consideration is a defence to restitution, it may be more accurately characterised as a bar to relief. Alternatively, good consideration serves to reverse an enrichment if invoked as a ground for counter-restitution.

informed by the common law.[94] More precisely, the focus must shift from overarching doctrine to the anatomy of the wrong. Equity views an agent's breach primarily as the betrayal of a relationship of trust and secondarily as a failure to observe the duty which underlies that relationship. From a common law perspective, the breach amounts to a failure of consideration. Enquiry is directed to construction of the contract and causation. If the wrongful conduct goes to the heart of the contract, it may be found that there was a failure of basis.[95] It is a question of degree. The contract will remain on foot and an agent may retain remuneration for his services if the breach of duty does not go to the whole contract,[96] or if the wrongdoing is merely incidental to the agent's performance,[97] an isolated event or severable from the core services contemplated by the agency agreement.[98]

In contrast, from an equitable perspective, forfeiture is unlikely to be granted if it would be disproportionate or inequitable.[99] As the language suggests, jurisdiction is exercised on a discretionary basis and the outcome is less predictable than equivalent proceedings at law. As Dyson LJ (as he then was) observed in declining an application for rescission:[100]

> When exercising its equitable jurisdiction the court considers what fairness requires not only when addressing the question of the precise form of relief, but also when considering whether the remedy should be granted at all.

The strict standards of fiduciary duty create a low threshold for breach,[101] but, once liability is established, equity is circumspect with regard to relief. The enquiry is normatively driven, focusing on the manner in which the fiduciary relationship has been abused[102] and its impact on the respective interests of the parties. While discretion is notionally exercised in a principled manner, the outcome is unpredictable. When a defaulting agent's contractual services have been performed and the principal has been fully compensated for any loss arising from a breach of duty, it is unclear whether an additional sanction of forfeiture will be imposed. This is both understandable and unsatisfactory. This state of affairs is not surprising given that an agency relationship is recognised both at law and in equity and judgments in this area reflect unreconciled elements of each. For this reason it is submitted that the common law's threshold of a failure of consideration would produce a more consistent and predictable outcome in equity and remove some of the more arcane distinctions between the two forms of action.

94 See A Burrows, 'We Do this at Common Law but that in Equity' (2002) 22 Oxford Journal of Legal Studies 1, where it is argued that substantive inconsistencies between common law and equity can sometimes be eliminated by a modest adjustment to their respective positions.

95 If there is a failure of basis, forfeiture can be granted as part of the process of rescission. An agent's commission is subject to forfeiture as the parties are restored to their pre-contract position. Thus, forfeiture is integral to *restitutio in integrum*. However, rescission assumes mutual, not unilateral obligations and counter-restitution can appropriately be invoked in relation to an executed agency contract.

96 In *Keppel v Wheeler* (n 12) 592, Atkin LJ accepted that 'there may well be breaches of duty which do not go to the whole contract, and which would not prevent the agent from recovering his remuneration'.

97 *Hippisley v Knee Brothers* [1905] 1 KB 1 (DC), 8.

98 *Peninsular & Oriental Steam Navigation Co v Johnson* (1938) 60 CLR 189, 216; *Nitedals Taendstikfabrik v Bruster* [1906] 2 Ch 671, 674–75; *Hippisley v Knee* (n 97) 9; *Erikson v Carr* (1945) 46 SR (NSW) 9 (Eq), 14–15; *Darmago v Fullerton Nominees Pty Ltd* 1999, SCWA CIV 1007, 27.

99 *Governor & Company of the Bank of Ireland v Jaffery* [2012] EWHC 1377 (Ch), [373]; *Avrahami v Biran* (n 89) [345].

100 *Johnson v EBS Pensioner Trustees Ltd* [2002] EWCA Civ 164, [79] per Dyson LJ.

101 See, for example, *Warman* (n 13)557.

102 The threshold is sometimes expressed in terms of whether the fiduciary acted in good faith or bad faith: *Maketu Estates Ltd v Robb* [2014] NZHC 2664.

It has been observed that the equitable standard for forfeiture is far from settled. The test is variously expressed as dishonesty, at one extreme, and equitable fraud, at the other. And traversing the remedial spectrum is an indeterminate test of unconscionability. In contrast, a single test of failure of consideration narrows the enquiry and promotes a more predictable outcome.

It may be objected that equitable discretion fosters uncertainty regardless of how the enquiry is framed. However, this can be overplayed. In fact, some argue that, as a body of principles, modern equity is 'as refined, rigorous and ultimately unyielding as anything produced by the common law'.[103] Moreover, it does not inevitably follow that common law reasoning in this area is more stable or immune from internal inconsistencies. For example, some judgments enunciate general principle without regard to the distinct status of secret commissions from a third party and agreed remuneration from a principal.[104] The latter gives rise to further uncertainties. The principal and agent will usually have defined the performance standards for the agent's services. Thus, the fundamental question whether there has been a total failure of consideration is inherently uncertain in that it is qualified by the nature of the agent's undertaking. However, the uncertainty is more factual than conceptual, and to that extent the distinct focus of failure of performance is a preferable model for both law and equity. In contrast to the current equitable approach to an agent's breach of fiduciary duty, a unitary test directed to the agent's performance provides clear direction as to the normative underpinnings of the enquiry. This is persuasively propounded by Peter Watts,[105] who argues that there is basically just one question: did the consideration fail due to the conduct of the agent? This leaves discretionary and unpredictable equitable considerations out of the equation and also removes the risk of punitive outcomes.[106]

6 Jurisdictional and doctrinal implications of a unitary standard

It has been argued that the equitable response to forfeiture of commission should be informed by common law principles. However, the challenges of adopting a unitary standard should not be understated. Even a modest assimilation of common law doctrine within the scheme of equitable wrongs raises significant policy[107] and taxonomic issues.[108] This draws to the fore the jurisdictional relationship between law and equity. A fundamental issue that continues to divide judges, practitioners and academics alike[109] is whether the Judicature Acts[110] effected a substantive fusion of the principles and doctrines of law and equity or whether each retains a separate and distinct identity. On one view, the Judicature Acts were procedural only and merely vested the administration

103 J Hackney, *Understanding Equity and Trusts* (Fontana Press 1987) 18–19. See also J Penner, *The Law of Trusts* (9th edn, Oxford University Press 2014) 4.

104 As previously noted, the former is a wrong per se, the latter is not.

105 Watts, 'Forfeiture of Agents' (n 4) 203.

106 See, for example, *Imageview* (n 5), discussed above.

107 See, for example, New Zealand's decision to accept incremental fusion of law and equity and its implications for determining the scope of fiduciary duty (*Day v Mead* [1987] 2 NZLR 443 (CA)).

108 The contemporary relevance of the debate is evident in *Fischer v Nemeske Pty Ltd* [2016] HCA 11, (2016) 330 ALR 1, where the majority of the High Court of Australia held that, with limited exceptions, a beneficiary's claim for money had and received will not lie against a trustee in the execution of her equitable obligations.

109 See, for example, the papers collected in S Degeling and J Edelman (eds), *Equity in Commercial Law* (Lawbook Co 2005).

110 The Supreme Court of Judicature Act 1873, ch 66, and the Supreme Court of Judicature Act 1875, ch 77.

of law and equity in one court.[111] As expressed in Ashburner's oft-quoted metaphor: 'the two streams of jurisdiction . . . though they run in the same channel, run side by side and do not mingle their waters'.[112] The alternative view is that equity and common law are effectively 'mingled or merged'.[113] The debate has profound ramifications in fashioning (or denying) jurisdictional boundaries and defining the nature and scope of obligations and remedies.[114] A *via media* between the competing views is to recognise that each body of law has, to some degree at least, retained distinct elements. This is to be expected given their separate development over many centuries. At the same time, neither equity nor common law operate in silos and it is unrealistic to suggest that each system is impervious to the influence of the other.[115] Before and after the passing of the Judicature Acts, both equity and common law had the capacity to adopt and adapt concepts from each other.[116] Developing this point, Sir Anthony Mason, writing extra-curially, opined:[117]

> The traditional principles of equity are not so invincibly superior to the concepts of the common law that equity cannot occasionally profit from common law ideas. And, though the courts should look at policy arguments with due circumspection, it would be absurd to suggest that the courts cannot adjust or modify equitable principle on policy grounds where to do so is appropriate.

Other commentators have gone further and suggested that the debate may have moved on and that the appropriate focus is no longer fusion or fusion fallacy, but the desirability of legal change and development of principle.[118]

A further qualification is necessary. Equity's assimilation of common law reasoning with respect to money had and received may be less radical than first appears. The doctrinal basis of an action for money had and received is commonly attributed to Lord Mansfield's judgment in *Moses v Macferlan*.[119] The case was heard in King's Bench, but the proceedings were said to be 'founded in the equity of the plaintiff's case'[120] and characterised as a 'kind of equitable action to recover back money, which ought not in justice to be kept'.[121] In similar vein, Lord Mansfield considered that the action was consistent 'with the judgment of the Court of Conscience'.[122] His Lordship subsequently described such proceedings as 'a liberal action in the nature of a bill in equity'.[123] The

111 *Salt v Cooper* (1880) 16 Ch D 544, 549. See further P V Baker, 'The Future of Equity' (1977) 93 Law Quarterly Review 529; J D Heydon, M J Leeming and P G Turner, *Meagher, Gummow and Lehane's Equity: Doctrines and Remedies* (5th edn, LexisNexis 2015) [2-110]–[2-150].

112 W Ashburner, *Ashburner's Principles of Equity* (Butterworths 1902) 23.

113 *Aquaculture Corporation v New Zealand Green Mussel Co* (n 37) 301 per Cooke P. See also *Day v Mead* (n 107) 451.

114 At times, discourse has been lively. For example, Lord Diplock's description of Ashburner's fluvial metaphor as 'mischievous and deceptive' (*United Scientific Holdings Ltd v Burnley Borough Council* [1978] AC 904 (HL), 924–25) was denounced by Meagher JA as being 'so erroneous as to be risible' (*GR Mailman & Associates v Wormold* (1991) 24 NSWLR 80 (NSWCA), 99).

115 *Harris v Digital Pulse* (n 37) [18]; *Norberg v Wynrib* [1992] 2 SCR 226, 272.

116 *Harris v Digital Pulse* (n 37) [141]. The interaction of law and equity with respect to money had and received is well documented. See, for example, *Muschinski v Dodds* (1985) 160 CLR 583, 619–20; *Baltic Shipping v Dillon* (n 79) 376; *Roxborough v Rothmans* (n 82) [100].

117 A Mason, 'The Place of Equity and Equitable Remedies in the Contemporary Common Law World' (1994) 110 Law Quarterly Review 238, 243–44.

118 Heydon et al (n 111) [2-380]–[2-400].

119 (1760) 2 Burr 1005, 97 ER 676 (KB). The influence of *Moses v Macferlan* is discussed in R Havelock, 'Rivalry over Liability for Defective Transfers' in Devonshire and Havelock (n 4) 119.

120 *Moses v Macferlan* (n 119) 1008 (Burr) and 678 (ER).

121 Ibid 1012 (Burr) and 680 (ER).

122 Ibid 1009 (Burr) and 678 (ER).

123 *Clarke v Shee* (1774) 1 Cowp 197 (KB), 199–200, 98 ER 1041, 1042.

elision of law and equity was exemplified by Lord Mansfield's openness to different sources of law, and was noted in Gummow J's influential judgment in *Roxborough v Rothmans of Pall Mall Australia Ltd*:[124]

> With varying degrees of success, Lord Mansfield sought to translate equitable principles, doctrines, and procedures into the trial of actions at law; this reflected his appreciation of equitable doctrine for its flexibility and adaptability to modern needs, particularly in commercial law.

His Honour elaborated that, following *Moses v Macferlan*, the action for money had and received has been recognised as equitable in character,[125] in England,[126] Australia[127] and the United States.[128] This was subsequently affirmed by the High Court of Australia in *Australian Financial Services & Leasing Pty Ltd v Hills Industries Ltd*.[129] The plurality observed that the equitable origin of the action for money had and received had long been recognised in Australia[130] and its provenance was undisputed:[131]

> There can be no denying the equitable roots of the principle by which a claim for restitution of money had and received to the use of the payer is to be determined. In *Dale v Sollet*, Lord Mansfield said of the action: 'This is an action for money had and received to the plaintiff's use. The plaintiff can recover no more than he is in conscience and equity entitled to' . . . In *Roxborough v Rothmans of Pall Mall Australia Ltd*, Gummow J explained that the 'equitable notions' of which Lord Mansfield wrote have been absorbed into the 'fabric of the common law' right of action for money had and received.

As Havelock aptly notes, money had and received may be regarded as an action that is equitable in nature, but which has historically been enforced by the common law.[132] Thus, equity's reception of a common law wrong does not necessitate the kind of doctrinal upheaval that might have been anticipated. In fact it is in line with the modest incremental developments that are generally accepted by both sides of the fusion/fusion fallacy debate.

7 Conclusion

The relationship of principal and agent is but one of a varied and indeterminate list of relationships that are characterised as fiduciary. Equity primarily regards breach of fiduciary duty as the betrayal of a relationship of trust. Comprehensive remedies can be invoked to purge the wrong and deter others from like conduct. Notwithstanding the strict standards that define fiduciary duty, it does not follow that the remedial response will be oppressive. Not uncommonly, equitable compensation places the principal in the same position as if the agent's duties had been faithfully performed. This may explain why in some cases equity compensates for loss but falls short of forfeiting an agent's

124 *Roxborough v Rothmans* (n 82) [84] (footnotes omitted).

125 Ibid [85]–[87].

126 *Kelly v Solari* (1841) 9 M & W 54, 58, 152 ER 24, 26; *Lodge v National Union Investment Co Ltd* [1907] 1 Ch 300, 311–12; *Royal Bank of Canada v King* [1913] AC 283 (PC), 296; *Westdeutsche Landesbank Girozentrale v Islington London Borough Council* [1994] 4 All ER 890 (QBD), 923.

127 *Campbell v Kitchen & Sons Ltd* (1910) 12 CLR 515, 531; *Muschinski v Dodds* (n 116) 619–20; *National Commercial Banking Corporation of Australia Ltd v Batty* (1986) 160 CLR 251, 268.

128 *Myers v Hurley Motor Co* (1927) 273 US 18, 24 (SC); *United States v Jefferson Electric Manufacturing Co* (1934) 291 US 386 (SC), 402–03; *Atlantic Coast Line Railroad Co v Florida* (1935) 295 US 301, 309.

129 [2014] HCA 14, (2014) 253 CLR 560.

130 Ibid [75].

131 Ibid [68]–[69] (footnotes omitted). See also [1], [65], [70], [71], [75] and [76].

132 R Havelock, 'Rivalry over Liability for Defective Transfers' in Devonshire and Havelock (n 4) 132, citing Gummow J's dictum in *Roxborough v Rothmans* (n 82).

commission. Whether the additional sanction of forfeiture should be imposed and the basis for doing so is problematic.

The forfeiture of an agent's commission is neither compensatory nor gain-based and is best regarded as a form of *sui generis* relief which is triggered by a breach of duty. The threshold is elusive. Equity does not countenance dishonesty or fraud in absolute terms. The interests of both parties are weighed in the scales, regardless of the gravity of the wrong. For example, an account of profits against dishonest and fraudulent fiduciaries[133] may be partly offset by allowances.[134] Again, an errant fiduciary may be granted an equitable lien or charge over trust property if the fiduciary has contributed value to that asset.[135] Equity is therefore circumspect before depriving an agent of remuneration if the principal has received the benefit of those services. Potentially, though, forfeiture can be granted across a spectrum of wrongs, from equitable fraud to cynical breach – although predictably, the more reprehensible the wrong, the more compelling the case for relief.

The common law is guided by the nature of the agent's undertaking and its contractual foundation. A claim for money had and received restores the *status quo ante* when performance has been frustrated by the agent's dishonesty. If the agent has repudiated the agreement and the basis on which his services were promised, the right to payment falls away. However, in some cases an agent may have performed the agreed contractual services, in whole or in part. Here, the interests of the parties can be adjusted to reflect the value of the agent's services. An agent's claim for counter-restitution, founded on good consideration, is a credible basis for that determination. This model should be applied in equity as well as law. An approach based on failure of consideration would produce a more consistent and principled outcome in equity. In this setting at least, equitable doctrine is not compromised by the reception of common law principles. On the contrary, as famously expressed in a different context, there are alloys which strengthen without corrupting.[136]

133 *Murad* (n 73).

134 *Crampton-Smith v Crampton-Smith* [2011] NZCA 308, [2012] 1 NZLR 5, [65]ff. See also *Chirnside v Fay* (n 56); *Boardman v Phipps* (n 13); *Warman* (n 13).

135 *Timber Engineering Co Pty Ltd v Anderson* [1980] 2 NSWLR 488, 499. See also *LAC Minerals Ltd v International Corona Resources Ltd* [1989] 2 SCR 574, (1989) 61 DLR (4th) 14.

136 P Devlin, *The Enforcement of Morals* (Oxford University Press 1965) 123.

NOTES AND COMMENTARIES

NILQ 70(2): 283–287

Stott, status and *stare decisis*

BRICE DICKSON
Queen's University Belfast

In November 2018 the UK Supreme Court issued an important decision on the discriminatory treatment of prisoners, *R (Stott) v Secretary of State for Justice*.[1] The issues involved were clearly difficult for the five Justices, since their judgments were not issued until more than 10 months after the appeal hearing had taken place, and in the end there were two prominent dissenters – the Court's President, Lady Hale, and its Deputy President, Lord Mance. The leading judgment for the majority was written by Lady Black, a relative newcomer to the court. In terms of paragraphs hers was the second longest judgment issued by any Supreme Court Justice during 2018 and the fifth longest issued since the Court was established in 2009.[2]

Mr Stott's claim was that he had been the victim of unlawful discrimination on the ground of his status because, as a prisoner serving an 'extended determinate sentence' (EDS) for a series of rapes,[3] he would become eligible for parole only after serving two-thirds of the 'appropriate custodial term',[4] which in his case was 21 years, whereas other categories of prisoners serving determinate sentences were eligible for parole after serving just one-half of their sentence. He claimed that this differential treatment amounted to a violation of Article 14 of the European Convention on Human Rights (ECHR), taken together with Article 5, which protects the right to liberty. Article 14 provides that: 'The enjoyment of the rights and freedoms set forth in this Convention shall be secured without discrimination on any ground such as sex, race, colour, language,

1 [2018] UKSC 59, [2018] 3 WLR 1831.

2 Lady Black's judgment was 156 paragraphs. In *In the matter of an application by the Northern Ireland Human Rights Commission for Judicial Review* [2018] UKSC 27, [2019] 1 All ER 173 Lord Kerr's judgment was 197 paragraphs. In *Lehman Brothers Holdings Inc v The Joint Administrators of Lehman Brothers International (Europe)* [2017] UKSC 38, [2018] AC 465 Lord Neuberger wrote a judgment of 187 paragraphs; in *Rubin v Eurofinance SA* [2012] UKSC 46, [2013] 1 AC 236 Lord Collins wrote 177 paragraphs; in *R (Walumba Lumba (Congo)) 1 and 2 v Secretary of State for the Home Department* [2011] UKSC 12, [2012] 1 AC 245 Lord Dyson wrote 169 paragraphs.

3 Imposed under the Criminal Justice Act 2003, s 226A, which was inserted by the Legal Aid, Sentencing and Punishment of Offenders Act 2012, s 124, and amended by the Offender Rehabilitation Act 2014, s 8(2). EDS replaced the much maligned 'indefinite sentence for public protection' (IPP); IPP prisoners were entitled to automatic release on licence after serving one-half of their custodial sentence.

4 Criminal Justice Act 2003, s 246A, inserted by the Legal Aid, Sentencing and Punishment of Offenders Act 2012, s 125(3), and amended by the Criminal Justice and Courts Act 2015, s 4(2) and (3).

religion, political or other opinion, national or social origin, association with a national minority, property, birth *or other status*' (emphasis added).

'Status' and precedent

At first instance the Divisional Court rejected the argument that the differential treatment was on the ground of Stott's 'status', but only because it felt constrained to do so by the doctrine of precedent. In *R (Clift) v Secretary of State for the Home Department*[5] the House of Lords had held that the different treatment of a prisoner serving a sentence of 15 years or more could not be said to be on the ground of his 'status'. As the Court of Appeal might have held that it too was bound by what the House of Lords had said in *Clift*, Stott was permitted to appeal directly to the Supreme Court under the leapfrog procedure.[6]

By four to one (Lord Carnwath dissenting), though without expressly referring to the 1966 Practice Statement governing the top court's approach to its own previous decisions,[7] the Supreme Court decided that the precedent in *Clift* should be set aside, largely because when the case was taken to Strasbourg the European Court of Human Rights (ECtHR) had ruled that Mr Clift *did* have a 'status' for the purposes of Article 14 of the ECHR and that he *had* suffered discrimination.[8] In *Stott* Lady Black helpfully summarised the respects in which the ECtHR had gone further than the House of Lords in *Clift*, and she examined three more recent decisions of the ECtHR, two of which also involved prisoners,[9] concluding that they confirmed the ECtHR's approach in *Clift*. Lady Black then analysed how the term 'status' had been dealt with by the House of Lords and Supreme Court in cases other than *Clift*. There were four such decisions prior to that case,[10] from which she extracted seven propositions. After looking at four further decisions issued after *Clift*[11] she remarked that just one of the seven propositions was no longer valid, namely the proposition that a person could have a 'status' only if he or she had a personal characteristic separate from the differential treatment being complained about. In *Clift* the ECtHR had made it clear that differential treatment could itself confer a 'status'.

This acceptance by the Supreme Court of a new approach to the concept of 'status' is significant because it greatly extends the potential of Article 14 to address instances of inequality. 'Status' no longer needs to be something analogous to the characteristics expressly mentioned in Article 14 as bases for discrimination claims. Not only is being a prisoner a 'status', being a particular type of prisoner can be a 'status' too. Lord Carnwath,

5 [2006] UKHL 54, [2007] 1 AC 484.

6 Administration of Justice Act 1969, ss 12–16. There are usually only two or three such appeals each year.

7 [1966] 1 WLR 1234. In *Austin v Southwark LBC* [2010] UKSC 28, [25] Lord Hope said, with the agreement of his four colleagues in that case, that the Practice Statement 'has as much effect in this Court as it did before the Appellate Committee in the House of Lords'.

8 *Clift v UK* App No 7205/07, judgment of 13 July 2010.

9 *Biao v Denmark* (2016) 64 EHRR 1; *Khamtokhu and Aksenchik v Russia* (GC) Apps Nos 60367/08 and 961/11, judgment of 24 January 2017; *Minter v UK* (2017) 65 EHRR SE6.

10 *R (S) v Chief Constable of the South Yorkshire Police* [2004] UKHL 39, [2004] 1 WLR 2196; *R (Hooper) v Secretary of State for Work and Pensions* [2005] UKHL 29, [2005] 1 WLR 1681; *AL (Serbia) v Secretary of State for the Home Department* [2008] UKHL 42, [2008] 1 WLR 1434; *R (RJM) v Secretary of State for Work and Pensions* [2008] UKHL 63, [2009] 1 AC 311.

11 *R (Kaiyam) v Secretary of State for Justice* [2014] UKSC 66, [2015] AC 1344; *Mathieson v Secretary of State for Work and Pensions* [2015] UKSC 47, [2015] 1 WLR 3250; *R (Tigere) v Secretary of State for Business, Innovation and Skills* [2015] UKSC 57, [2015] 1 WLR 3820, *R v Docherty (Shaun)* [2016] UKSC 62, [2017] 1 WLR 181.

who is known for being a rather conservative judge,[12] did not think that the ECtHR had gone as far as the majority thought it had in this area. Citing *Minter v UK*,[13] where the ECtHR decided an application was inadmissible because it was based on the flimsy argument that prisoners sentenced under a different regime from one used earlier were treated preferentially, he stated:

> I would need considerable persuasion that the authors of the Convention intended mere conviction of a criminal offence, or subjection to a particular custodial regime, to entitle the recipient to specially protected status under human rights law. More generally, it is important that Article 14 is kept within its proper role within the Convention, and outside the core protected areas is not allowed to develop into a means of bypassing the carefully defined limits applicable to the individual rights.[14]

Fortunately the majority in *Stott* clearly rejected Lord Carnwath's position, but the case as a whole provides further evidence of how careful the Supreme Court now is to locate its human rights decisions within the jurisprudence of the ECtHR. None of the Justices treated the position of the ECtHR as binding, especially as there was no Grand Chamber judgment on the point, but they all accorded it deep respect and tried to be consistent with it. In *Poshteh v Royal Borough of Kensington and Chelsea* the Supreme Court said that before it would change a domestic precedent in the light of the jurisprudence of the ECtHR it would await a full consideration of the matter by the Grand Chamber of that court.[15] However, as we know from the more recent decision in *R (Hallam) v Secretary of State for Justice*,[16] even when there is a Grand Chamber judgment in play Supreme Court Justices may still disagree about its import and may therefore remain reluctant to depart from one of their own precedents.

Analogous situations and justification

Having held in favour of Mr Stott on the status point, the court then found by three to two (Lady Hale and Lord Mance dissenting) that on the facts before them the discrimination in question was justified. In accordance with custom and practice all the judges dealt with the justification issue in two parts, even though the wording of Article 14 does not explicitly require such an approach. They first considered whether Mr Stott was in an 'analogous situation' to other prisoners who were treated differently. On this the majority's view was summed up by Lady Black in this fashion:

12 He was one of three dissenters in the Supreme Court's decision in the Brexit case, *R (Miller) v Secretary of State for Exiting the EU* [2017] UKSC 5, [2018] AC 61. Since 2015 he has also manifested a traditionalist streak when dissenting in *Airtours Holidays Transport Ltd v Commissioners for HM's Revenue and Customs* [2016] UKSC 21, [2016] 4 WLR 87, on the meaning of 'supplying services' for the purposes of VAT law; *Impact Funding Solutions Ltd v Barrington Support Services Ltd* [2016] UKSC 57, [2017] AC 73, on the interpretation of exemption clauses in insurance contracts; *Goldtrail Travel Ltd (In Liquidation) v Onur Air Tasimacilik AS* [2017] UKSC 57, [2017] 1 WLR 3014, on making an appeal conditional upon a payment into court; and *Regency Villas Title Ltd v Diamond Resorts (Europe) Ltd* [2018] UKSC 57, [2018] 3 WLR 1603, on the reach of the law on easements.

13 See (n 9) above. *Minter* refers also to *Massey v UK* App No 14399/02, judgment of 8 April 2003, an application made by a prisoner who lost in the High Court: see *R (Massey) v Secretary of State for Justice* [2013] EWHC 1950 (Admin).

14 See (n 1) [179].

15 [2017] UKSC 36, [2017] AC 624.

16 [2019] UKSC 2, [2019] 2 WLR 440.

> . . . the various sentencing regimes have to be viewed as whole entities, each with its own particular, different, mix of ingredients, designed for a particular set of circumstances.[17]

That ruling would have been enough to reject Stott's appeal on the basis that he was not in an analogous situation to that of other prisoners, but the majority added that, even if it was in an analogous situation, the difference in how he was treated as far as eligibility for parole was concerned was objectively justifiable. They observed that EDS prisoners were different from other prisoners with determinate sentences because the former *must* first have been designated as presenting a significant risk of serious harm to members of the public.[18] Moreover, compared with prisoners serving indeterminate sentences, EDS prisoners were in at least one respect in a better position because they had a guaranteed end date to their incarceration.[19]

Lady Hale and Lord Mance thought that there was no justification for requiring EDS prisoners to remain in prison for two-thirds of the custodial term appropriate to the seriousness of their offending while discretionary life sentence prisoners were eligible for release after just one-half of what would have been the appropriate determinate sentence for their conduct. Lady Hale pointed out that a discretionary life sentence prisoner is even more dangerous than an EDS prisoner,[20] although Lord Hodge said that that was not necessarily the case.[21] She stressed that the most important question any prisoner wants an answer to is 'When will I get out?'

In truth, the requirement that a discrimination claimant must show that he or she is in an analogous situation to someone who is treated more advantageously seems to be an unnecessary addition to the requirement that any difference in treatment be justifiable. The latter should surely embrace the former, in the sense that if there is an objectively justifiable reason why two prisoners are treated differently it ought to be clearly explicable by those who are applying the treatment in the first place. Merely saying that the prisoners have been sentenced under different regimes should not be enough. The erection of additional hurdles which alleged victims of inequality have to overcome if they are to succeed in a discrimination claim is tantamount to reinforcing the systemic nature of some forms of inequality. Notoriously, systemic discrimination is the hardest kind of discrimination to combat because systems are often the product of ingrained assumptions and bureaucratic complacency. It is obvious that the sentencing of convicted criminals is a complicated process, one which inevitably requires a wide range of personal factors to be taken into account, but the process should not be made additionally complex by creating further differences between criminals based supposedly, but often spuriously, on penological grounds. The average onlooker, not to mention the average prisoner, would undoubtedly detect a fundamental unfairness in such an approach, especially if the difference manifests itself in the calculation of when the prisoner becomes eligible for release.

Although it was not an issue in *Stott*, UK discrimination law also suffers from maintaining the distinction between direct and indirect discrimination, a distinction which the ECtHR manages to do without. In recent years the UK Supreme Court has often been divided over whether particular forms of discrimination fall into the direct or

17 See (n 1) [155]. At [193] Lord Hodge agreed: 'It is appropriate to take a holistic approach to each sentencing regime in deciding whether or not one regime is analogous to another.'

18 Ibid [146]. See the Criminal Justice Act 2003, s 226A(1)(b).

19 Ibid [155].

20 Ibid [218].

21 Ibid [193].

indirect category, most noticeably in *R (E) v Governing Body of JFS*,[22] *Bull v Hall*[23] and *R (SG) v Secretary of State for Work and Pensions*.[24] One reason for the distinction is meant to be that there are some forms of discrimination which should never be justifiable (direct discrimination) and others which should occasionally be justifiable (indirect discrimination), but as well as introducing a demarcation line which is difficult to define this ignores the fact that in extreme cases even so-called direct discrimination is and should be permitted.[25]

When deciding whether discriminatory enjoyment of a right is justifiable the nature of the right in question surely matters. In *Stott* Lady Black conceded that because the right to liberty is such an important right differential early release schemes need to be carefully scrutinised lest they violate it.[26] If Messrs Clift or Stott had been complaining not about their additional loss of liberty but about their lesser entitlement to visits, enhanced 'pay' for prison labour or 'single cell occupancy', the majority's calculus as to whether such adverse treatment was justifiable may well have been different. In the assessment of justification more intense scrutiny is bound to be given to differential treatment affecting a person's liberty than to such treatment affecting the right to a private or family life. At the same time, although the right to liberty is not one of the qualified rights in the ECHR, it is not one of the absolute rights either: some prisoners may therefore enjoy a lesser right to liberty than others, as the outcome to the *Stott* case indicates. But it would be better if the differential right to liberty were expressed in the length of sentence imposed on prisoners in the first place, not in the rules on their eligibility for parole. If the explanation for requiring some prisoners to serve two-thirds of their sentence before being eligible for parole is public protection and public confidence in criminal sentencing,[27] those prisoners should be given a longer sentence in the first place. In *Stott* the majority unfortunately accepted, as the ECtHR has done, that courts should interfere with penal policy only if it renders the right to liberty arbitrary – a high threshold.[28]

For countries that have ratified Protocol 12 to the ECHR, which extends the Article 14 right not to be discriminated against in relation to enjoyment of 'the rights and freedoms set forth in this Convention' to enjoyment of 'any right set forth by law', there will be many other opportunities for claims to be made based on status, and there too justifiability will often be the key issue. Regrettably neither the UK nor Ireland has ratified Protocol 12, and there seems no immediate prospect of their doing so.[29]

22 [2009] UKSC 15, [2010] 2 AC 728.

23 [2013] UKSC 73, [2013] 1 WLR 3741.

24 [2015] UKSC 16, [2015] 1 WLR 1449.

25 An employer can discriminate against an employee directly on the basis of his or her age if the employer can show that this was a proportionate way of achieving a legitimate aim: Equality Act 2010, s 13(2). Moreover, non-disabled persons and single persons cannot claim discrimination on grounds of disability or marital status.

26 See (n 1) [81]. On this she cited the warning of the ECtHR in *Clift v UK* (n 8) above, paras 62 and 73.

27 These are the reasons given by Lady Black at [152]–[154] and Lord Hodge at [199].

28 See in particular Lord Hodge at [198]–[200], referring to *Clift v UK* (n 8) above, paras 73 and 74.

29 Twenty countries have so far done so, including Finland, the Netherlands and Portugal.