JOURNAL OF MEDIA LAW
https://doi.org/10.1080/17577632.2023.2216523

Routledge
Taylor & Francis Group

Interpretation *is* opinion: realigning the fact/opinion distinction in English defamation law

Thomas DC Bennett

City Law School, City, University of London, London, UK

ABSTRACT

Statements that interpret the words of others are – by their nature – interpretative. Their meaning is contingent and (inter)subjective; it is something that those who encounter those words or actions construct. As such, it makes little sense for the law to treat interpretative statements as purporting to set out provable facts. Yet English defamation law does precisely this. As a result, publicly criticising ambiguous words or actions that are put into the public domain by one person can cost a commentator dearly. This essay critiques the peculiarities of English defamation doctrine that have created this situation, arguing that all published statements that interpret or purport to interpret one or more earlier statements, whether expressly or impliedly, ought to be treated as statements of opinion. By adopting this approach, English defamation law can address an issue that has the potential to cause significant chilling effects on public discourse – particularly on social media.

KEYWORDS Defamation; libel; fact/opinion distinction; honest opinion; truth

Introduction

A novel type of defamation case has appeared on the legal landscape in England and Wales, owing principally to the emergence of social media – and, particularly, Twitter – as a key medium for political communication and debate. This case type involves the claimant making an ambiguous statement which provokes opinionated and perhaps vituperative public criticism from the defendant, following which the claimant sues the defendant for libel. The early judicial response to this sort of case, illustrated by the decision of the High Court in *Riley v Murray*,[1] is deeply problematic. For it indicates that statements

CONTACT Thomas DC Bennett thomas.bennett.2@city.ac.uk

[1] *Riley v Murray* [2021] EWHC 3437 (QB), [2022] EMLR 8, [8] ('trial judgment').

This article has been corrected with minor changes. These changes do not impact the academic content of the article.

which interpret earlier statements are at high risk of being labelled assertions of fact, rather than expressions of opinion. They can avoid this only by adhering to several exacting legal requirements of which the vast majority of social media users will be wholly unaware. This is problematic for two reasons. First, *any* statement that interprets (explicitly or tacitly) an earlier statement can *only* properly be regarded as an expression of opinion. This is because interpretation is a reader-led process that is necessarily subjective. Second, treating such statements in this way has the potential to lead to serious chilling effects on public discourse, by drawing those exercised by another's words but ignorant of the law's exacting requirements into something of a trap.

The argument presented in this essay is that, going forward, English libel law should treat statements which interpret earlier statements as being, *prima facie*, assertions of opinion. This should form a general rule. In order to make this argument, I will draw on 'reader response' theory and, in particular, Stanley Fish's concept of 'interpretative communities'. These, Fish tells us, constitute the space within which the interpretation of texts occurs and the meaning of those texts comes into being.[2] Reader response theory and Fish's 'interpretative communities' give us tools with which we can make sense of what is happening when the law grapples with the fact/opinion distinction in these sorts of cases.

I will focus the argument around a detailed critique of the reasoning in *Riley* – in particular of the reasoning of the preliminary hearing judgment at which the issue of the offending statement's meaning was considered – since that case is exemplary of a problem that, if not resolved, we are most likely to see many more examples of in the near future.[3]

Riley v Murray

In order to bring this matter into focus, we will first examine the exemplar case of *Riley v Murray*, which illustrates perfectly the problem with which we are concerned.

Riley relates to two 'tweets' sent on the social media platform, Twitter. The first tweet was posted by the claimant, the well-known TV presenter, Rachel Riley, in March 2019. Riley's tweet was itself a 'retweet' (a copy, republished, in this case with some additional content added by Riley) of an earlier tweet by the political commentator Owen Jones. Two months earlier, in January of the same year, Jones had posted his original tweet after an egg had been

[2]See Stanley Fish, *Is There a Text in This Class?* (Harvard University Press 1980), and *Doing What Comes Naturally* (OUP 1989). Fish actually calls these 'interpretive communities', preferring the more popular spelling ('intepretive') in his native America over the more common spelling in the UK ('interpretative'). Rather than switch between the two, I will exclusively use the British version, 'interpretative', and will adjust Fish's original text where quoted to reflect this. No further indication of such adjustments will be given.

[3]*Riley v Murray* [2020] EWHC 977 (QB) ('meaning judgment').

thrown at the far-right political figure and former leader of the British National Party, Nick Griffin. In his tweet, Jones opined: 'sound life advice is, if you don't want eggs thrown at you, don't be a Nazi.' Riley's March tweet, posted an hour after the then leader of the Labour party, Jeremy Corbyn, had had an egg thrown at him during a visit to a mosque, added to Jones' tweet her own words – 'Good advice.' – and emoji (character-sized pictures selected from a range of available emoji on the platform) of a red rose and an egg. It is well-known that the Labour party's emblem is a red rose. The judge in this case, Mr Justice Nicklin, found as a matter of fact that Riley 'was referring to the incident involving the assault on Mr Corbyn earlier that afternoon'.[4] It is contextually relevant and well-documented that Riley has publicly criticised both Jeremy Corbyn and the Labour party over his and its records on tackling anti-Semitism in the party.[5]

If the claimant in this libel action had been Jeremy Corbyn, suing Riley for insinuating that he was a Nazi, or someone who held such Nazi-like views that he deserved public violence, the case would take on a much more familiar form.[6] But, in this case, Riley was the claimant, and the defendant was a woman called Laura Murray. Murray, who was one of Corbyn's aides, saw Riley's tweet and interpreted it as legitimising what she saw as a violent

[4] *Riley*, (trial judgment) (n 1).
[5] ibid [2].
[6] It would, for example, bear some resemblance to the infamous tweet in *McAlpine v Bercow* [2013] EWHC 1342 (QB).

attack on Corbyn. Later the same evening, Murray posted her own tweet to her own Twitter profile (page), in which she said:

> Today Jeremy Corbyn went to his local mosque for Visit My Mosque Day, and was attacked by a Brexiteer. Rachel Riley tweets that Corbyn deserves to be violently attacked because he is a Nazi. This woman is as dangerous as she is stupid. Nobody should engage with her. Ever.

This tweet was the catalyst for the litigation in this case, as Riley sued Murray on the basis that it was defamatory of her. On the date she published the tweet, Murray had 7,245 followers on Twitter. The tweet was 'retweeted' (copied and reposted by other users) 1,544 times and 'liked' (given a positive upvote) by 4,738 people.

At a preliminary hearing to determine the meaning borne by the offending statement, Riley pleaded that Murray's tweet, in its ordinary and natural meaning, meant:

> The Claimant had publicly supported a violent attack upon Jeremy Corbyn at a mosque by saying that he deserved it. She has shown herself to be a dangerous person who incites unlawful violence and thuggery and is therefore so beyond the pale that people should boycott her and her tweets.[7]

Murray pleaded that the statement bore a different meaning, and that the tweet, as a whole, was an expression of opinion.[8] This argument was broadly along the lines of that which this essay argues is the most appropriate way to deal with this sort of case. It was summarised by Nicklin J in the following terms:

> Mr Hudson QC submits that the hypothetical ordinary reasonable reader: "*would understand [the Tweet] to mean that the Claimant had posted a tweet (following the attack on Jeremy Corbyn) which (in the Defendant's view) meant (or conveyed) that Jeremy Corbyn deserves to be violently attacked because he is a Nazi*".[9]

Nicklin J, however, found that the natural and ordinary meaning of the statement was:

(1) Jeremy Corbyn had been attacked when he visited a mosque.
(2) The Claimant had publicly stated in a tweet that he deserved to be violently attacked.
(3) By so doing, the Claimant has shown herself to be a dangerous and stupid person who risked inciting unlawful violence. People should not engage with her.[10]

[7] *Riley* (meaning judgment), (n 3) [2].
[8] ibid [8].
[9] ibid [22] (emphasis is original).
[10] ibid [25].

He clarified further that '[p]aragraphs (1) and (2) are statements of fact. Paragraph (3) is an expression of opinion.'[11] This is where the problem with which we are concerned arises. For Nicklin J finds that the words 'Rachel Riley tweets that Corbyn deserves to be violently attacked because he is a Nazi' amount to an assertion of fact. As we will see, it is precisely these words that should be regarded as an expression of opinion. In order to contextualise our analysis, we must first turn our attention to the 'single meaning rule'.

The 'single meaning rule'

English defamation law contains an idiosyncratic rule known as the 'single meaning rule' (SMR). According to this rule, the court must identify the one and only meaning of any statement alleged to be defamatory. When juries were still used to determine questions of fact, the judge and jury each played separate parts in the SMR's operation: the judge excluded unreasonable meanings (thereby pre-empting perverse findings from the jury), but the jury otherwise had a free hand to determine the single meaning which they found the statement to bear. Since the Defamation Act 2013 effectively abolished jury trial, however, the operation of the SMR has been the sole province of the judge. The SMR inquiry takes place, typically, at a preliminary hearing. At this stage, the parties make submissions on the issue of meaning, and the judge decides what the single meaning of the statement is, which delimits the range of possible defences for trial.

The SMR has been criticised by judges and scholars alike on the basis that it has little grounding in reality. The objections to it are twofold. First, statements invariably bear multiple (perhaps even infinite) meanings (for reasons we will shortly explore) and simply do not bear a single, objectively-determinable meaning. Second, even if this were not the case and statements did in fact bear only a single meaning, judges are ill-positioned to identify this meaning.

Lord Nicholls called the SMR 'a crude yardstick', since readers 'vary enormously in the way they read articles and the way they interpret what they read.'[12] In *Ajinomoto*, Sedley LJ thought the rule 'anomalous, frequently otiose and, where not otiose, unjust';[13] 'a pragmatic practice [that] became elevated into a rule of law and has remained in place without any enduring rationale.'[14] Rimer LJ piled on in the same case in even stronger terms, saying that if the SMR managed to achieve a fair balance between the parties' interests, it 'would appear to be the result of luck rather than judgment'.[15]

[11] ibid [25].
[12] *Charleston v NGN Ltd* [1995] 2 AC 65, 73–74.
[13] *Ajinomoto Sweeteners v Asda Stores Ltd* [2010] EWCA Civ 609; [2011] QB 497 [31]
[14] ibid [32] (Ajinomoto).
[15] ibid.

The leading practitioner text on defamation, *Gatley on Libel and Slander*, is similarly critical of the rule, observing that

> [b]y insisting on a single meaning when it is clear that the words may carry more than one meaning, the court will either fail to provide redress for injury that has unquestionably been suffered or overcompensate the claimant by awarding him damages for a meaning that some readers may have found the words to bear whereas others, wholly reasonably, understood the words in a non-defamatory or less-defamatory sense.[16]

Perhaps the most obvious problem with the SMR when it comes to the *Riley* case is that it gives rise to a particularly peculiar situation. Riley's tweet is found to be ambiguous and thereby capable of being interpreted in multiple ways.[17] Indeed, Murray loses the case at trial because her own tweet did not explicitly state that Riley's tweet could be interpreted in ways other than that in which she (Murray) interpreted it.[18] Riley's statement, not being the statement complained of in the case, was not subject to the SMR, yet Murray's own statement was. The fact that it, too, could be interpreted in a range of ways, was not *legally relevant to the judge's decision*, since defamation doctrine requires the statement complained of to be attributed just one, single meaning. This indicates that English defamation law in this field lacks conceptual coherence. For the courts readily and simultaneously accept: (a) that statements have multiple meanings, and (b) that statements have only a single meaning. Moreover, the courts distinguish between these positions not on the basis of how language works, but rather on what the requirements of the relevant doctrine are. In other words, the court adopts an obviously incoherent position simply on the basis that the doctrine obliges it to do so, with occasional riffs on the idea that this is a 'pragmatic'[19] position designed to enable litigation to move forward, which seem more apologetic than justificatory.[20]

In order to throw light onto the issue of the SMR, we can usefully draw (at a level of some abstraction) on the discipline known as literary criticism. Literary criticism is concerned with reading and interpreting texts. As such, locating the root of a text's meaning is a core concern for literary critics. Within this field, there has been a long-running debate as to where a text's meaning comes from. Formalists – those whose analyses 'approach the [text's] immanent or internal architecture' – in literary criticism (as in law) locate the root of meaning in the text itself.[21] But for others, meaning

[16] Richard Parkes and Godwin Busuttil eds, *Gatley on Libel and Slander* (13th edn, Sweet & Maxwell 2022), 97–98.
[17] *Riley* (trial judgment), (n 1), [74], [157].
[18] ibid [77]–[78], [124]–[126], [130]–[131].
[19] In the lay, rather than philosophical, sense of being outcome-oriented.
[20] See *Stocker v Stocker* [2019] UKSC 17, [2020] AC 593, [34].
[21] Jim Hansen, 'Formalism and Its Malcontents: Benjamin and De Man on the Function of Allegory', (2004) 35(4) New Literary History 663, 666.

is constructed by the reader. This second school of thought is broadly known as 'reader response theory'. It encompasses several variants (one of which we will return to in more detail later), but the broad thrust of the theory is that meaning is subjective; it arises through the act of interpreting a text undertaken by its reader.[22]

This is not the place in which to rehearse the large and complex argument between formalists and proponents of reader response. Fortunately, it is not necessary to do so, for one, simple, doctrinal reason: English defamation law *already operates* on the express basis that meaning is constructed by those to whom defamatory statements are published. In other words, defamation law presumes the correctness of a premise that is *also* foundational in reader response theory. This is the case because, in English defamation law, it matters not a jot whether the author intended a defamatory meaning; it matters only that a hypothetical reader interprets the statement as defamatory.[23] English law rejects the notion that the text itself is the only relevant source of data for this question through recognising the potential for 'innuendo' meanings.[24] (This is in spite of an express rule that, when determining the meaning of a statement, only the statement itself can be taken into account.[25] The courts resolve this tension by drawing a distinction between additional information located in other texts (which is excluded), and information likely to be already known to the reader (which is included).[26]) Thus, a foundational premise of defamation law – that which instructs us on how a defamatory meaning comes into being – is based on the notion that meaning is constructed by the reader (or audience) in their mind, through the act of interpretation.

A corollary of this, which is perhaps made most explicit via the innuendo rule, is that the meaning of a defamatory statement may (and almost certainly will) be different for each reader. For when readers interpret statements, they do so bringing their own experience to bear. For example, the statement 'she has left me' will likely mean something different to an adult than it does to a child; an adult, with experience of romantic relationships,

[22]For a broad introduction to this field of scholarship, see Jane P Tompkins (ed), *Reader-Response Criticism: From Formalism to Post-Structuralism* (John Hopkins University Press 1980).

[23]*Bolton v O'Brien* (1885) 16 LR Ir 97, 108 (May CJ) and 118 (O'Brien J).

[24]*Lewis v Daily Telegraph* [1964] AC 234, 278–9 (Lord Devlin). A legal 'innuendo' is a compound meaning whereby a *prima facie* inoffensive statement is coupled, in the reader's mind, with some pre-existing, extrinsic knowledge the reader possesses, resulting in a novel meaning that is defamatory. For example: Andrea tells Sameerah that Bert is sleeping with Jenny. Sameerah knows that Bert is married to Minnie. Thus, in Sameerah's mind, Andrea has told her that Bert is being unfaithful to Minnie.

[25]*Telnikoff v Matusevitch* [1992] 2 AC 343.

[26]And, yes, this causes a further tension between the hypothetical standard of the 'ordinary, reasonable reader', and the presence of particular actual readers who actually possess the requisite knowledge to form an innuendo meaning. This is a tension for which defamation law appears to have no principled justification.

will likely take this as indicating the break-up of such a relationship, whilst a child is likely to take the statement more literally, in the sense of being left behind somewhere. This bears relation to linguistic philosopher Ludwig Wittgenstein's notion of 'aspect perception' and its opposite, 'aspect blindness'; Wittgenstein tells us we will only be able to interpret what we see by reference to our past experience.[27]

A problem arises, however, for defamation law when the SMR comes into contact with the reality that different readers will interpret the same statement in entirely separate, subjective ways, thereby attributing to it a uniquely personal meaning. At the heart of defamation law, then, is a paradox: the statement must, by law, have one single meaning; but since meaning is acknowledged to be constructed by the statement's readers, it *cannot* have a *single* meaning. The law's solution is to *impose* upon the statement a single meaning. This it does by the application of a test which aims, explicitly, to identify – on a balance of probabilities – the single meaning that most likely reflects the meaning that a hypothetical character known as the 'ordinary, reasonable reader' would attribute to the statement.[28] This tactic is common in tort law. In negligence, for example, the standard of care is set by reference to another hypothetical character – the 'reasonable man'.[29]

Here, my critique of the law's operation could proceed down one of two paths. I could make the argument that, since the test is invented, interpreted and applied by judges, the standard of the 'ordinary, reasonable reader' is a judicial construct that inevitably fails to reflect in any meaningful way real individuals out there in the world and that, therefore, this supposedly 'objective' test simply reflects the experience and internal biases of the judicial class. This would be a highly sceptical argument, according to which the SMR is a wholly artificial norm that removes reality from the equation entirely. Lord Nicholls' critique of the SMR in *Bonnick*, though less strident, tends in this direction,[30] as does the criticism in *Gatley*.[31]

Alternatively, I could take a less sceptical path. If we accept that judges are making genuine and at least partially successful efforts to imagine the perspective of an 'ordinary, reasonable reader', then the test aims not to identify *the* single meaning, but to select *a* single, *most likely* meaning. This approach might be defended as pragmatic (in the lay, rather than philosophical, sense), since it focuses on making the legal rules in this field workable; by identifying

[27] See Ludwig Wittgenstein, *Philosophical Investigations* (3rd edn, GEM Anscombe tr GEM Anscombe and P Rhees eds, Basil Blackwell 1978) 194ff (on 'aspect perception') and 213–4 (on 'aspect blindness'), For a useful discussion of 'aspect perception' and 'aspect blindness' in the tort context, see Emilia Mickiewicz, 'An Exploratory Theory of Legal Coherence in Canengus and Beyond' (2010) 7 The Journal Jurisprudence 465.

[28] Although, as I say, this creates a tension with the innuendo rule itself in principle. See (n 26).

[29] *Hall v Brooklands Auto Racing* [1933] 1 KB 205, 224 (Greer LJ).

[30] See *Bonnick v Morris* [2002] UKPC 31; [2003] 1 A.C. 300, [21] and [24] (Lord Nicholls).

[31] See *Gatley*, (n 16).

the meaning around which it is most likely that the average readership would coalesce, the litigation is able to proceed.[32] Were this not the case, defenders of this approach would evince, English tort law would struggle to get purchase on the case at hand, since it is not equipped to grapple with multiple, concurrent meanings.

Both the sceptical and less sceptical readings of what judges in this field actually do in practice are entirely plausible. This is hardly surprising. Since those of us who analyse judicial activity and legal norms in this field are interpreting what we see, we do so in a range of ways that reflect our individual experience; sceptics and judges will interpret *the same actions* in different ways.

The sceptical path, however, leads only in a direction with which many legal analysts are already familiar. It is a classically realist path, which will point up (what realists see as) the inherent tendency of judging, as a practice, to reflect internal judicial biases, and will (likely) end up offering a blanket criticism of the SMR as arbitrary and divorced from the real world it seeks to govern.[33] It is not a path that is likely to secure a great deal of attention from those it critiques, who may interpret it as an attack on *judges* rather than *judging* (plausible interpretations, both). So, for the sake of trying to keep a larger audience on board, so to speak, I will focus on the less sceptical path. I will assume that judges are trying to locate a meaning that has some basis in the minds of others, rather than simply their own (thereby replicating the historic role of the jury), and that they might even have some success in this regard. Unfortunately, even if they achieve this, the approach remains problematic – so much so that it needs to change. This can be seen when we consider the matter through a framework given to us by the literary critic Stanley Fish in his work on 'interpretative communities'.

'Interpretative communities'

Stanley Fish's work on 'interpretative communities' is at once highly influential and highly controversial. The idea aims to explain how meaning, constructed in the minds of individual readers, can nonetheless be shared by numerous individuals exposed to the same text. In other words, it aims to explain how a whole bunch of different people can interpret a text in a sufficiently similar way that they can get together and discuss it, act upon it, critique it, and so forth. For if meaning is entirely subjective, how would language-based communication even be possible?

[32]This is very much *not* philosophical pragmatism. Pragmatic philosophers would be appalled.
[33]See generally EW Thomas, *The Judicial Process: Realism, Pragmatism, Practical Reasoning and Principles* (CUP 2005).

Fish's answer to this is that readers are not isolated creatures. They exist within what he calls 'interpretative communities'; it is these communities that, acting with some sort of group-think, give rise to shared meanings through the interpretation of texts. These communities are made up of individuals with shared background experiences. In particular, interpretative communities are

> made up of those who share interpretative strategies not for reading (in the conventional sense) but for writing texts, for constituting their properties and assigning their intentions.[34]

Moreover,

> if the understanding of the people in question [is] informed by the same notions of what counts as a fact, of what is central, peripheral, and worthy of being noticed – in short, by the same interpretative principles – the agreement between them will be assured, and its source will not be a text that enforces its own perception but a way of perceiving that results on the emergence to those who share it.[35]

Fish himself tells us that what he means by 'interpretative community' is well-captured by Gerald Postema in a piece of writing focusing on the work of Ronald Dworkin. Postema tells us that

> [t]o learn a social practice is to become acquainted through participation with a new common world; it is to enter a world already constituted. One does not bring an understanding (let alone a theory) to the practice; rather, through participation one comes to grasp, tentatively and uncertainly at first, then more securely, then critically, the common meaning of the practice.[36]

Such a community, or practice, 'takes the form of a shared discipline and a thick continuity of experience of the common world of the practice'.[37] What Fish is describing, then, is a community within which there are common understandings of certain concepts and ideas, which are picked up and internalised by those who embed themselves within that community. In this way, the language used by members of an interpretative community exhibits the sort of 'agreement in judgments' that Wittgenstein tells us is essential for linguistic communication.[38] As a member of the community, individuals will play a role in developing the shared understandings within the community, but cannot unilaterally alter them for the community.

[34]Fish, 'Interpreting the Variorum' in *Is There a Text in this Class?* (n 2), 171.
[35]Fish, *Doing What Comes Naturally*, (n 2), 153.
[36]Gerald J Postema, '"Protestant" Interpretation and Social Practices' (1987) 6 Law and Philosophy 283, 313 (quoted in Fish, *Doing What Comes Naturally*, (n 2), 580).
[37]ibid 318–19.
[38]Wittgenstein, (n 27), 88 (§242).

From the outset it must be acknowledged that Fish's theory has been subjected to a great deal of criticism. One criticism is that his definition(s) of 'interpretative communities' are so vague and abstract as to make discussion of them difficult.[39] As a result, it could be said (cynically) that any attempt to criticise what Fish means by 'interpretative communities' might be written off as not reflective of what Fish meant, giving Fish a seemingly (and nominatively appropriate) watertight defence. Moreover, anyone making use of his theory can use his highly malleable definition(s) to support a wide range of understandings of just what such a community looks like, so an interpretative community could plausibly take on any shape that such a scholar needed it to.[40] Fish's theory could thus be said to be rather slippery.

I therefore do not seek to claim that what I present here is anything other than my understanding – my interpretation – of Fish's theory. My understanding is, of course, constructed by my experience – an experience shared by the 'interpretative community' (or, indeed, communities) of which I am a part (for instance, the community of legal scholars). As with reader response theory more generally, however, our task is made simpler by the fact that English defamation law already recognises not only the existence, but the relevance to the task of determining meaning, of a range of interpretative communities.

English defamation law uses the hypothetical 'ordinary, reasonable reader' as a theoretically objective standard by which to determine the meaning of a statement.[41] This character has a number of characteristics that, like the absolutely-average-in-every-conceivable-respect 'reasonable man' in negligence,[42] ultimately render him, her or them a rather unrealistic character. This ordinary, reasonable reader is thus said to be

> not naïve but … not unduly suspicious. He can read between the lines. He can read in an implication more readily than a lawyer and may indulge in a certain amount of loose thinking but he must be treated as being a man who is not avid for scandal and someone who does not, and should not, select one bad meaning where other non-defamatory meanings are available.[43]

[39] Erie Martha Roberts 'Something Fishy is Going On: The Misapplication of Interpretive Communities in Literary Theory' (2006) 1(1) The Delta, Article 6, 2.

[40] Ibid.

[41] Sceptics would, of course, aver that *any* standard applied by judges is inevitably coloured by those judges' personal, unconscious biases. For, as the American realist jurist and scholar Benjamin N Cardozo observed, judges are no more able to escape that instinctive 'stream of tendency' within each of us 'which gives coherence and direction to thought and action' than other mortals. See *The Nature of the Judicial Process* (Yale University Press 1991) 12.

[42] *Hall*, (n 29).

[43] *Koutsogiannis v Random House* [2019] EWHC 48 (QB), [2020] 4 WLR 25, [12]. Nicklin J uses gendered language in this paragraph, which aligns with similarly gendered language used historically by the courts in setting out rules of tort law (such as negligence's 'reasonable man', per *Hall v Brooklands*, (n 29)). I have chosen not to do so in my own discussion of the 'reasonable reader' in the hope that the courts, in applying this judicial guidance, will not do so in a substantively gender-biased fashion. It is nevertheless noteworthy that English defamation law has long had this pointedly gendered aspect, the impact of which is beyond both my expertise and the space afforded in this essay, but which ought to be recognised.

However, the ordinary, reasonable reader is not taken to be representative of society as a whole. Instead, he/she/they is/are taken to be the ordinary, reasonable reader *of the particular publication complained of*. Thus, if the alleged libel appears in *The Times*, the ordinary, reasonable reader is taken to be the ordinary, reasonable *Times* reader.[44] If the statement appears on Facebook, they are the ordinary, reasonable Facebook user.[45] In this way, it is absolutely clear that, at least formally, English law recognises distinct interpretative communities for different outlets in which statements are published.

The leading authority on the application of the SMR in respect of statements published on social media is the Supreme Court judgment in *Stocker v Stocker*.[46] In this (unanimous) decision, the Supreme Court effectively confirms that, where a statement is published on social media, the 'ordinary, reasonable reader' test should be modified to take account of the fact that the statement's readers will mostly comprise other social media users. In other words, the Supreme Court both recognises the existence of an interpretative community of social media users, and declares that, in applying the 'ordinary, reasonable reader' test, the judges must seek to 'inhabit the world' of such a reader.[47]

Stocker concerned an allegedly defamatory post published by the defendant on the social media platform Facebook, in which the defendant had stated that the claimant, her former husband, had 'tried to strangle' her. In construing the meaning most likely to be attached to the statement by an ordinary, reasonable reader, the Supreme Court held, the judge ought to have taken into account interpretative norms peculiar to users of social media (and, more particularly, Facebook). The Court was absolutely explicit about the importance of this, stating '[t]he fact that this was a Facebook post is critical.'[48] When considering how the reasonable Facebook reader (as we might call this character) would interpret a post on the platform, the Court tells us that the test 'should reflect the circumstance that this is a casual medium; it is in the nature of conversation rather than carefully chosen expression; and that it is pre-eminently one in which the reader reads and passes on.'[49]

> [Facebook users] do not pause and reflect. They do not ponder on what meaning the statement might possibly bear. Their reaction to the post is impressionistic and fleeting.[50]

The core of the instruction being given by the Supreme Court here is that Facebook users do not typically engage in highly technical or close reading of the posts they read; rather, they take a quick look, get the gist of the

[44] *John v Times Newspapers Ltd* [2012] EWHC 2751 (QB), [19].
[45] *Stocker*, (n 20), [38].
[46] ibid.
[47] ibid [38].
[48] ibid [41].
[49] ibid [43].
[50] ibid [44].

post, and move along swiftly to another post. As such, the reasonable Facebook reader is likely to give a meaning to the statement that reflects the general gist of it, rather than specifics.

There are several cases, in addition to *Riley* (some of which in fact precede the Supreme Court's judgment in *Stocker*, and which the Supreme Court mentions with approval), in which the High Court has attempted to apply this same notion to statements 'tweeted' on Twitter. In *Monir v Wood*, Nicklin J himself said:

> Largely, the meaning that an ordinary reasonable reader will receive from a Tweet is likely to be more impressionistic than, say, from a newspaper article … The essential message that is being conveyed by a Tweet is likely to be absorbed quickly by the reader.[51]

Perhaps the most detailed guidance on applying the SMR in the context of tweets on Twitter is that given by Warby J in *Monroe v Hopkins*, when he observed that Twitter 'is a conversational medium[] so it would be wrong to engage in elaborate analysis of a 140 character tweet' and that 'an impressionistic approach is much more fitting and appropriate to the medium'.[52]

What these pieces of guidance broadly indicate is that English defamation law has already adopted something akin to Fish's notion of interpretative communities. In determining the single meaning of a statement, the 'ordinary, reasonable reader' test takes into account the likely readership of the particular medium in which the statement appears, and invites judges to attempt to imagine the characteristics peculiar to that readership.

Having established two important things – that English defamation law embraces the same foundational principle as 'reader response theory', and also recognises something akin to 'interpretative communities' – we are in a position to make the argument that statements of the sort made by Murray in *Riley* ought henceforth to be recognised and treated as statements of opinion. Support for this argument can be found in both reader response theory and existing doctrine. In the sections that follow, we will consider each of these in turn.

Support for my argument in reader response theory

Twitter and Facebook

Stocker concerned publication on Facebook, rather than Twitter. Whilst both are social media platforms, they function differently and have different communities of users, engaging with the platforms through different norms of interaction. For starters, the quantity of users in the United Kingdom differs significantly. The UK's population currently comprises approximately

[51] *Monir v Wood* [2018] EWHC (QB) 3525, [90].
[52] *Monroe v Hopkins* [2017] EWHC 433 (QB); [2017] 4 WLR 68, [35].

67.2 m people.[53] There are 55.8 m Facebook users based in the UK – indicating that the vast majority of the UK's population (statistically encompassing every UK resident between the ages of 14–80) has a Facebook account.[54] By contrast, there are just 18.4 m UK-based Twitter accounts, a figure representing just under one in four UK residents.[55]

Despite it being arguable that, statistically speaking, the average person in the UK – the 'reasonable person' – *is* a Facebook user, the Supreme Court in *Stocker* nonetheless chose to identify a distinct interpretative community of Facebook users, bringing its own distinctive experience to bear on the act of interpreting statements. Clearly – explicitly – the court thought this was necessary. Given that, again, speaking in quantitative terms, the average person in the UK, whilst a Facebook user, is *not* a Twitter user, it is surely clear that Twitter users must also be regarded as constituting a distinct interpretative community. Indeed, statistically speaking, we might expect members of the Twitter community to be *more* distinct from the average UK resident in terms of their characteristics, since that community is several times smaller than the Facebook community and, as we will see, the two platforms differ significantly in their functionality.

A key difference between Facebook and Twitter is the character limit that Twitter puts on individual tweets (once 140 characters, now 280 characters).[56] Despite creator Jack Dorsey's original intention that Twitter should be a place where users would share 'inconsequential' snippets of information,[57] it has become one of the most important tools for political expression and discussion in the world. This is perhaps epitomised by the infamous use of the platform by the 45th President of the United States, which eventually culminated in his being banned by the platform for breaches of its terms of use. Twitter's character limit causes users to find creative ways of sharing all of the information they wish in a highly condensed fashion. Strategies adopted by users include: using non-standard grammar, omitting punctuation, using contractions and acronyms, using emoji, omitting superfluous words such as articles and conjunctions, and

[53] https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/articles/overviewoftheukpopulation/2020 (accessed 25/4/23).

[54] https://www.statista.com/statistics/1012080/uk-monthly-numbers-facebook-users/ (accessed 25/4/23). There are 10.1m children under 14 in the UK, and 2.7m adults over 80. This totals 12.8m across the two age brackets, which is the difference between the total UK population and the number of UK-based Facebook users. (https://www.statista.com/statistics/281208/population-of-the-england-by-age-group/, accessed 25/4/23). Statistically, then, every UK resident between the ages of 14 and 80 is a Facebook user (excluding business accounts, which I have not been able to quantify).

[55] https://www.statista.com/statistics/242606/number-of-active-twitter-users-in-selected-countries/ (accessed 25/4/23).

[56] 'Characters' in this sense means typed characters, and includes letters, numbers, punctuation marks and spaces.

[57] David Sarno, 'Twitter creator Jack Dorsey illuminates the site's founding document. Part I' *Los Angeles Times*, 18 February 2009. https://www.latimes.com/archives/blogs/technology-blog/story/2009-02-18/twitter-creator-jack-dorsey-illuminates-the-sites-founding-document-part-i (accessed 25/4/23).

omitting segments of information deemed less relevant. Twitter users across the UK (and, indeed, globally – at least when using English) will recognise these strategies. They are unique (in their confluence, at least) to this interpretative community, and are embedded in its interpretative practice to the extent that most users will not even think twice about them when reading or writing tweets.

Twitter is also a medium in which users, by virtue of its ease of access and the short format of tweets, are tacitly encouraged to post content regularly and without a great deal of detailed consideration. If ever there was a social media format designed to elicit streams of consciousness from its users, Twitter is it. It is also noteworthy that tweets, once 'tweeted' (posted), cannot be edited (unlike Facebook posts).[58] Errors thus cannot be corrected. The only option for a user who wishes to alter the content of a tweet is to delete it and post a replacement. Again, this design feature is not a matter of accident. It appears designed to discourage users from dwelling on what they have already posted and instead to move on and post more, new content.

Expressing opinions on Twitter: interpretations and value judgments

Twitter's 'reply' function enables users to post tweets directly beneath (and thereby understood as relating to) others' tweets. This process creates a 'thread' based on the original tweet, and it is in these threads that discussions of the original tweet's content often unfolds. Replies are also limited to 280 characters. Several things result from this. One is that tweets and replies on matters of controversy tend to be expressed in short, often blunt, statements that frequently adopt a declarative format.

The courts have no real difficulty in identifying expressions of opinion that take the form of declarative statements. As is the case in other fields of law, substance rightly takes precedence over form. We see this in cases like *Butt*, *Greenstein* and *Ware*.[59] If a declarative statement is, at its core, a value judgment, this renders it a statement of opinion.[60]

The crucial difference in the *Riley* case appears to be that the defendant, Murray, failed to include within her tweet either a direct reference or a 'hyperlink' (a clickable weblink) to Riley's original tweet. Nicklin J repeatedly characterises Murray's tweet as having 'misrepresented'[61] Riley's by (a) presenting a summary of Riley's tweet that presented readers with just one

[58]Following Twitter's takeover by Elon Musk, there appear to be plans to allow some, limited editing of tweets for certain categories of user, but there is no indication that this will become universal.
[59]*Butt v Secretary of State for the Home Department* [2017] EWHC 2619 (QB); *Greenstein v Campaign Against Antisemitism* [2019] EWHC 281 (QB); *Ware v Wimborne-Idrissi* [2021] EWHC 2296 (QB).
[60]*British Chiropractic Association v Singh* [2010] EWCA Civ 350; [2011] 1 WLR 133.
[61]*Riley* (trial judgment) (n 1), at [77], [124], [130] and [131].

interpretation (Murray's) of it (when Riley's tweet was in fact ambiguous), and (b) failing to include a hyperlink so that readers could locate Riley's tweet instantly and see its ambiguity for themselves.[62] As such, the 'ordinary, reasonable reader', according to Nicklin J, would not have sufficient knowledge of the Riley tweet for it to play a role in contextualising Murray's. That being the case, Murray's declarative tweet was considered a bald assertion, and was therefore treated as an assertion of fact. This is Nicklin J's conclusion, from the precedents laid before him in argument and/or that he was already aware of, as to how English defamation doctrine applies to the facts of the case.

We will return in detail to these precedents and the decision to apply them in this case in the next section. For now, let us accept the fact of their application and explore reasons to believe those precedents have underestimated the ability of the ordinary, reasonable Twitter user – the hypothetical, average member of this interpretative community – to distinguish an assertion of fact from an expression of opinion in the absence of any reference to underlying fact.

The second sentence of Murray's tweet – 'Rachel Riley tweets that Jeremy Corbyn deserves to be violently attacked' – does not appear to contain a value judgment. But it contains something *akin* to a value judgment in its effect: an explicit reference to the fact of the statement being an interpretation of an earlier statement. Murray's use of the word 'tweets' here is judicious. It is a usefully concise (given the character limit) way of indicating that the particular statement of Riley's upon which her (Murray's) interpretation is based is contained in a tweet which can be found on Riley's feed.[63] The reasonable Twitter user will pick up on this. Moreover, Murray's use of the present tense – 'tweets' – indicates that the particular post to which she is referring is broadly contemporaneous with her own. Thus, the reasonable Twitter user will understand that they need not trawl back through days of posts on Riley's feed, but will locate the relevant tweet within broadly the same time frame as Murray's. Nicklin J does not appear to have attributed this sort of contextual knowledge to the reasonable Twitter user. Rather, he takes the view that, unless Murray's tweet either contained a hyperlink to Riley's original post, or was posted as a direct reply on Riley's feed, the reasonable Twitter user would not be able to locate the relevant underlying facts (Riley's tweet). This is unrealistic.

[62] See particularly ibid [77]–[78].

[63] This is something the English courts have attributed significance to before. In *Telnikoff v Matusevitch* [1992] 2 AC 343, the House of Lords acknowledged that placing part of a statement inside quotation marks would lead the ordinary, reasonable reader to conclude that the quoted portion was something that the person to whom the quote was attributed had actually said. It is reasonable to assume that the ordinary, reasonable reader would draw the reverse conclusion in the absence of quotation marks.

First, hyperlinks on Twitter are often lengthy since they involve including the entire URL in the tweet, which takes up a substantial number of characters (quite probably enough to put Murray's tweet over the limit and render it impossible to post without cutting other parts of its content). Second, locating Riley's Twitter feed independently is not akin to going and finding an earlier edition of a newspaper, or some other document. It is more akin to turning to another page in the same newspaper or book. This is something which the courts have consistently held to be part of the relevant context that the ordinary, reasonable reader is taken to have read.[64] For Twitter is a platform upon which locating individual feeds is quick and easy. Riley's feed, in particular, is easy to locate since, by virtue of her high number of 'followers' (subscribers), simply typing 'Rachel R' into the search box brings her profile (@RachelRileyRR) up as a prompt which can be accessed with a single click. This is, without serious doubt, far simpler than manually scouring other pages in a newspaper or book for facts supporting an opinion featuring elsewhere in the same publication – an act which the courts routinely presume the reader to have undertaken.[65]

What, then, makes an explicit interpretation 'akin' to a value judgment? Reader response theory provides us with an answer. According to any variant of reader response theory (whether Fishian or otherwise), texts are rendered meaningful only through the act of interpretation. *Any* interpretation of another's words – including the original author's interpretation of their own words – is necessarily *subjective*. If Fish's theory of interpretative communities is correct, then interpretations are *inter*subjective. But the broadly subjective nature of the act of interpretation holds on any model of meaning based on reader response. Because defamation law's internal recognition that readers bring their own interpretation to texts is what leads defamation to create the hypothetical 'ordinary, reasonable reader' character in the first place, it stands to reason that one of the characteristics of that character must be a self-awareness that the meaning of a statement is constructed in their own mind. If this were not the case, defamation law would contain a bizarre paradox wherein a hypothetical character who exists *only* because the law accepts as foundational a particular theory of meaning was not themselves presumed to acknowledge the very theory that brings them into (their hypothetical state of) being. Thus, we can presume that the ordinary, reasonable reader themselves recognises that a later statement which purports to interpret an earlier statement is, by rendering an interpretation of that earlier statement (which is then itself interpreted by the ordinary, reasonable reader of the later statement), saying something that is the subjective view of the (later) author. Put simply, if I state (without

[64] *Chalmers v Payne* (1835) 2 Cr M&R 156, 159; *Charleston v NGN Ltd* [1995] 2 AC 65, 73–74.
[65] ibid.

indicating quotation) that 'Tom says X', the ordinary, reasonable reader of my statement must be taken to have interpreted 'X' as being my subjective interpretation of what Tom said, not an objective, factual account of Tom's utterances.[66]

That this is the case is, surely, an inescapable component of ordinary communication. Consider the work of academics. Regularly we make assertions of this sort, where we say that Ronald Dworkin says X, or the Supreme Court says Y. When we do so, it is clear to the ordinary, reasonable reader of our work (who exists, presumably, in an interpretative community of readers of legal scholarship) that we are giving our interpretation of what Dworkin or the Court said. This *is* our opinion.

Imagine that I write of Ronald Dworkin that his work portrays a theory of law that broadly fits the description 'natural law'. Some scholars would agree with this statement.[67] Many others would disagree.[68] Ronald Dworkin himself (were he still with us) might conceivably take offence at my characterisation of his work. But none of these responses would alter the nature of my statement: what I have said is necessarily my interpretation of his work and thus necessarily a matter of my (inter)subjective opinion. Now imagine further that, like Laura Murray, I have said nothing more about Dworkin's work in support of my bald statement. I have not quoted him, nor linked to a website where his publications might be read. I have not even indicated which of Dworkin's many publications have informed my opinion. Is my statement any less of an opinion for these deficiencies? No. I may have difficulty *justifying* my opinion, or in persuading a court that I have provided sufficient information to indicate to an ordinary, reasonable reader the basis of the opinion to satisfy the requirements of s.3 Defamation Act 2013. But the fact I may not be able to *defend* my opinion does not prevent my statement from *being* an opinion. Yet if my statement is ruled to be an assertion of fact, the defence I am left with is to prove my statement to be substantially true under s.2.[69] That is something that is impossible. I cannot prove the truth of my opinion, because my opinion is not amenable to proof. I cannot prove that my interpretation of Dworkin's work is the one and only true interpretation, because it is inescapable that there will be other, rival interpretations. (Even if Dworkin were still alive, his own interpretation of his own work would be just that – an interpretation.)

However, the decision in *Riley v Murray* requires me (in principle) either to present the reader with an easy way of reading the relevant parts of

[66]The ordinary, reasonable reader would, of course, likely come to a different conclusion if I obviously quoted Tom, by saying 'Tom says "X"'.

[67]See eg Stephen R Perry, 'Judicial Obligation, Precedent and the Common Law' (1987) 7(2) OJLS 215, 216; David AJ Richards, 'Taking *Taking Rights Seriously* Seriously: Reflections on Dworkin and the American Revival of Natural Law' (1977) 52(6) NYU Law Review 1265.

[68]Eg Postema, (n 36).

[69]Or subject to the new form of public interest privilege under s.4, under which I would still need to have taken steps to verify the truth of what I had said prior to publication.

Dworkin's work (which copyright laws probably prevent me from doing, at least in the online context), or to *expressly* acknowledge *in the course of my statement* that, to parody an old BBC phrase, 'other interpretations are available'. Of course other interpretations are available! An ordinary, reasonable Twitter user who recognises – as they will – what I am doing as giving my interpretation will also *necessarily* realise that my opinion is not the only one that could be held.

Support for my argument in defamation doctrine

In the 'meaning' judgment in *Riley*, Nicklin J finds that Murray's statement that 'Rachel Riley tweets that Jeremy Corbyn deserves to be violently attacked because he is a Nazi' is an assertion of fact. I have made an argument that the nature of this statement is necessarily one of opinion because, as an interpretation of another's (Riley's) earlier statement, it can only be the expression of Murray's (inter)subjective understanding of the earlier text. The reasonable Twitter user, moreover, will not fail to recognise this, given their familiarity with the formatting constraints of the platform and the norms of robust debate that take place there. Even if my argument on this point is persuasive, however, it *prima facie* encounters a doctrinal difficulty that must be dealt with.

The Kemsley rule

The difficulty is that the courts have a long-standing practice of treating certain statements as assertions of fact even if they are not. The classic example was given by Lord Porter in the 1952 House of Lords case of *Kemsley v Foot*, in which the common law defence of 'fair comment' was in issue. Lord Porter quotes approvingly from what was at the time a leading text on defamation, *Odgers on Libel and Slander*:

> If the defendant accurately states what some public man has really done, and then asserts that 'such conduct is disgraceful,' this is merely the expression of his opinion, his comment on the plaintiff's conduct. … But if he asserts that the plaintiff has been guilty of disgraceful conduct, and does not state what that conduct was, this is an allegation of fact for which there is no defence but privilege or truth.[70]

The language here rather obscures a matter that requires some unpacking. For *Odgers* says that, in the second scenario, 'this *is* an allegation of fact', rather than 'this *will be treated as* an allegation of fact'. But it is plain to see what such a statement really is: it is an unsupported expression of opinion – a 'bald assertion' (to use a phrase often ascribed to this sort of statement in the defamation literature).

[70] *Kemsley v Foot* [1952] AC 345, 356, quoting from *Odgers on Libel and Slander* (6th edn, 1929), at 166.

The more modern discussion of this rule, however (which I will call the '*Kemsley* rule'), of which the leading example (cited with approval by the Court of Appeal since[71]) is from one of Nicklin J's own judgments, accepts that such statements are 'by their nature' opinion, but that they will nevertheless be 'treated as statements of fact':

> Some statements which are, by their nature and appearance opinion, are nevertheless treated as statements of fact where, for instance, the opinion implies that a claimant has done something but does not indicate what that something is, i.e. the statement is a bare comment.[72]

So, the old common law on fair comment contained a rule whereby unsupported expressions of opinion would be *treated as* assertions of fact. There was no hard-and-fast threshold for determining the extent of support required for an opinion to be treated as an opinion, rather than a fact. Rather, the courts saw this as a matter of context, ultimately answerable only by reference to what the ordinary, reasonable reader would be in a position to do. If the offending statement provided the reader with sufficient information about the underlying facts for the reader to be in a position to judge for themselves whether the opinion was a 'fair' one, the statement would be treated as one of opinion. But if it did not give the reader enough information, it would be treated as a statement of fact in itself. This old common law rule continues to be deployed today (including in *Riley*), despite s.3 Defamation Act 2013 having abolished the fair comment defence and replaced it with a statutory defence of 'honest opinion'. The significance of this legislative change is considered in the next subsection; for now, we simply note the continued application of the *Kemsley* rule.

There is also another long-standing rule in defamation that the defendant is unable to pray in aid any extraneous factual information beyond the offending publication in seeking to persuade the court of the statement's meaning.[73] But the enforcement of this rule is not absolutely strict. Where the offending statement is a review of some matter placed by the claimant into the public domain – say a novel, a play, or a restaurant – it is obviously evident that the review will not usually contain the actual material upon which the opinion is based. The best the defendant can do is to give their own summary of the material.[74] The courts have applied the rule in such cases more loosely, counting the summary as the inclusion of relevant facts. But such quasi-factual summaries will, of course, be coloured by the defendant's opinion of the material under review, and so it is in no sense the sort of objective factual matrix that this

[71] See *Millett v Corbyn* [2021] EWCA Civ 567; [2021] EMLR 19, [12] (Warby LJ).
[72] *Kousogiannis v The Random House Group Ltd* [2019] EWHC 48 (QB); [2020] 4 WLR 25, [12] (Nicklin J).
[73] See eg *Telnikoff v Matusevitch* [1992] 2 AC 343, 352 (Lord Keith). Also *Kousogiannis*, ibid.
[74] See eg *Burstein v Associated Newspapers Ltd* [2007] EWCA Civ 600.

rule – which aims to provide readers with sufficient information to judge the fairness of the opinion *for themselves*, objectively – seems to envisage.[75]

The relevance of this is that the *Riley* case is far more like a review than a spontaneous personal attack. Riley put something into the public domain – her original tweet – which is then ripe for comment and criticism. Murray criticised it, and also criticised Riley for tweeting it. Yet the court did not give Murray the benefit of the looser application of the extraneous facts rule found in review cases, holding that Murray's summary of Riley's tweet, and Murray's indication that what Riley had said was locatable in a tweet on Riley's Twitter feed (which readers could find and read for themselves), was insufficient to qualify her (Murray's) statement to be treated as an opinion.[76]

'Honest opinion'

The foregoing brings us to s.3 Defamation Act 2013, which provides for the defence of 'honest opinion'. My argument here is that the courts have, in recent years, erred in their interpretation of this section which, on its proper construction, does not require *and does not permit* the courts to take into account the presence or absence of supporting facts in an offending statement when making a *preliminary* ruling as to whether the statement is one of fact or opinion.

We have seen that the older common law on 'fair comment' – the *Kemsley* rule – takes the presence or absence of supporting facts into account when making such a determination. However, this practice needs to be seen in its context. Prior to the 2013 Act, jury trials were common (although not ubiquitous) in defamation cases in England. When jury trials occurred, there was no need for a preliminary hearing to determine whether the offending statement was one of fact or opinion. It was for the jury to determine whether they felt the statement contained sufficient underlying facts to enable an ordinary, reasonable reader to decide for themselves whether the statement was one of opinion and, if it was, whether that opinion was 'fair'. Crucially, the two issues of whether the statement was one of fact or opinion *and* whether the opinion was 'fair' were decided together, *at the same stage in proceedings*.

With the effective abolition of jury trial by the 2013 Act, modern practice sees these two issues determined *separately*: the first at a preliminary hearing, and the second (reformulated to ask whether the opinion is one that an honest person in

[75]For example, a review might criticise an actor's performance for being 'tedious', citing as its factual basis the 'fact' that he 'lumbers slowly about the stage with all the elegance of an arthritic wasp'. The 'fact' is clearly coloured by the low opinion the writer has of the performance. But what the review in this sort of scenario does do is *point to the source material*, such that the reader could (theoretically) go and take a look at it *for themselves*. A food critic might write a ruinously bad review of a restaurant and paint a supposedly factual picture of it that is hideous, but so long as he identifies the restaurant, the reader could go and dine there themselves, and thereby decide whether his review was a fair one.

[76]*Riley* (meaning judgment) n 3, [28].

possession of the facts could hold) at trial. When considering the first question at the preliminary hearing, however, the doctrine to which the courts turn for guidance is invariably precedent from the pre-2013 era. It is entirely defensible that pre-2013 Act case law conflated the issues of whether the statement was one of fact or opinion *and* whether the opinion was 'fair', since – at that time – the two issues were dealt with simultaneously. But uncritically applying these precedents to post-2013 cases is problematic – not least because it runs counter to the wording of the statute and, seemingly, to Parliament's intention.

Section 3 provides:

(1) It is a defence to an action for defamation for the defendant to show that the following conditions are met.
(2) The first condition is that the statement complained of was a statement of opinion.
(3) The second condition is that the statement complained of indicated, whether in general or specific terms, the basis of the opinion.
(4) The third condition is that an honest person could have held the opinion on the basis of—
    (a) any fact which existed at the time the statement complained of was published;
    (b) anything asserted to be a fact in a privileged statement published before the statement complained of.
(5) The defence is defeated if the claimant shows that the defendant did not hold the opinion.

Note that subsections 2 and 3 expressly separate the issues of whether the offending statement is one of fact or opinion, and whether the offending statement indicates the basis of the opinion. These are clearly labelled and enumerated as separate 'conditions' by the statute. Indeed, the second condition (subsection 3) *presupposes* a finding that the first condition has been met *and therefore that the offending statement is one of opinion*. Thus, on a literal reading of the statute, the second condition is not part of the first. The implication of this is crystal clear: whether a statement is an assertion of fact or an expression of opinion does not turn on whether that statement indicates the basis of the opinion, and is presumed already to have been decided separately from that issue.[77] Thus far, however, no court has expressly considered this particular point in this level of detail.[78]

[77]The explanatory notes to the Defamation Act 2013 do not support the continued use of the *Kemsley* rule either. They are not binding, but they support the reading I have outlined. For they, too, enumerate these conditions separately.

[78]In *Butt* (n 59), the Court of Appeal discussed broadly (at [30]–[40]) the applicability of s.3 to the fact-opinion distinction at the preliminary stage, but did not consider the particular points I have made. The Court simply said that 'there was no discernable [sic] difference between the parties as to what [the] principles were' in respect of that distinction (at [33]), and took from the Explanatory Notes to the 2013 Act only the broadest

The older precedents on the *Kemsley* rule are of no real assistance. This is because they relate to the common law defence of 'fair comment' which is expressly abolished by s.3(8). As such, they provide no binding precedent on the novel statutory defence of 'honest opinion'. At best, those precedents can provide a degree of non-binding guidance, but that guidance cannot be followed if it conflicts with express provisions in the statute. Clearly, the *Kemsley* rule conflicts with the approach set out in detail in s.3. For it conflates two issues which are expressly enumerated as separate conditions by the statute. As every first-year law student knows, that which is expressly included in one part of a statute is impliedly excluded from other parts. So, whether on a literal reading or by operation of the *expressio unis est exclusio alterius* maxim, the *Kemsley* rule ought no longer to be being followed.

Moreover, even if I am wrong and ss.3(2) and 3(3) Defamation Act 2013 do not have the effect of negativing the *Kemsley* rule, s.3(3) expressly provides that the defendant need only indicate 'in *general or* specific' terms the basis for the opinion. Murray's use of the word 'tweets' meets, I would suggest, the requirement to indicate the basis of the opinion in 'general' terms, for the reasons outlined above. The court having found otherwise suggests an overly prescriptive interpretation of what 'general terms' requires, which is far more stringent than that applied in other review-type cases.

A better way

In light of the foregoing analysis, I propose a new general rule that would, if adopted, provide a solution to the key problems outlined in this essay whilst adhering to the core doctrinal principles of post-2013 defamation law. The rule proposed is that any statement that expressly or impliedly interprets an earlier statement is to be regarded as a statement of opinion. As such, the appropriate defence for such a statement (in respect of which a *prima facie* case for liability in defamation can be made out) would be the s.3 defence of 'honest opinion'.

The generality of the proposed rule reflects the broad, implied consensus in defamation law that reader response theory governs the determination of meaning in this field. It would improve the way that cases revolving around posts on Twitter and other social media are dealt with by resolving the problems identified in the *Riley* case. But it would, more broadly, improve coherence in this field by ensuring that the central notion of reader response as the authoritative determinant of meaning is not diluted by hazy judicial notions of what the experiences of particular interpretative communities are. In this way, the proposed rule can 'future-proof' defamation law in this burgeoning age of vibrant online debate.

statement in [21] that 'Condition 1 … is intended to reflect the current law'. The point I have made pertaining to the clear separation between conditions 1 and 2 (ss.3(2) and 3(3)) was not addressed.

It is true that the proposed rule is broad. So broad, in fact, as to encompass a wide range of statements previously considered, without a second thought, to be assertions of fact. Take, for example, the notorious case of *Irving v Penguin Books Ltd*.[79] David Irving is a Holocaust-denying anti-Semite who wrote several books about Nazi Germany. An American historian, Deborah Lipstadt, wrote a book in which she took both Irving and his work to task for its historical inaccuracies which, she claimed, stemmed from the way that his racist views coloured his interpretation of historical facts. Irving sued Lipstadt and her publisher, claiming that her labelling of him as a Holocaust-denying anti-Semite was defamatory. At trial, Lipstadt's defence of justification[80] succeeded on the basis that she successfully proved the allegations that she made to be substantially true. Under the proposed rule, however, the case would unfold quite differently. Lipstadt's allegations would not be regarded as assertions of fact at all, but expressions of her opinion. This is because her allegations are, by necessity, the result of her interpretation of Irving's work, which he has put into the public domain voluntarily. Her interpretation is one that will, no doubt, be shared by many – and most defensibly so. But reader response theory tells us that this fact – that the meaning she attributes to Irving's work arises through her act of interpreting it – is inescapable. So the appropriate defence today would not be truth but honest opinion.

This state of affairs does not seem particularly objectionable, in normative terms. Indeed, the courts have already found allegations of anti-Semitism and racism generally to be expressions of opinion in other cases.[81] In the *Irving* example, in order to make out the opinion defence, Lipstadt would have to show that Irving's work contains qualities that could lead an honest person to conclude that he is a Holocaust-denying anti-Semite, and that she has indicated – in general or specific terms – that Irving's work is the source material from which she has formed this opinion. This would not be difficult, given the substance of the material Irving has put into the public domain. Taking the approach I propose would thus allow for freer public debate of controversial statements made publicly by individuals.

There might be some instinctive objection to one particular implication of my argument, which is that, in a case like Lipstadt's, we might want to be able to say of her book that it speaks the 'truth' about Irving. There is emotional force to this, for there is understandable desire to label Irving's misreporting of history as empirically false. But maintaining that what Lipstadt is doing is asserting facts rather than her opinion has the counterproductive effect of making her defence far more difficult than it needs or ought to be. If we truly want to assist those in Lipstadt's situation going forward, my approach

[79] *Irving v Penguin Books Ltd* [2000] EWHC QB 115; [2000] 4 WLUK 339; 2000 WL 362478 (QB).
[80] The common law defence of justification was abolished and replaced, in the Defamation Act 2013, by the defence of truth under s.2.
[81] See eg *Greenstein v Campaign Against Antisemitism* [2019] EWHC 281 (QB); *Miller v Turner* [2021] EWHC 2135 (QB).

is preferable. For it is likely generally to be easier to make out the defence of honest opinion than that of truth. Nonetheless, it is no free pass; defendants must still show that there is a sufficient factual matrix underpinning the opinion to enable an honest person to hold that opinion. In this way, there is still plenty of scope for claimants to protect their reputation against unfounded, reputationally-damaging interpretations of their statements.

The benefits of freeing ordinary citizens from the potentially ruinous cost of libel litigation for spontaneous social media posts that interpret earlier ones ought to be clear. For one thing, there is a clear risk of chilling effects on public discourse if the approach taken in *Riley* becomes the new normal. Posts of the sort Murray tweeted – unsupported or barely-supported, declaratively-formatted and often blunt assertions of opinion – are commonplace, particularly on Twitter. Most troubling are those cases like *Riley* in which the claimant has posted a statement that is provocative (intentionally or otherwise) and open to multiple interpretations (which all statements are, but some more instinctively obviously than others), but then takes umbrage at one or more interpretative responses to that statement. Such a claimant might be thought to be wanting to have their cake and eat it too; she wants to say something controversial whilst insulating herself from criticism. In such cases, defamation scholars may well incline to the view that an opinion-based defence ought at least to be *available*, even if, ultimately, the defendant is unable to make out its elements. That is, claimants in these circumstances ought to have a valid claim if the opinion is not supported by facts or honestly held, but ought otherwise to tolerate the criticism they have invited by placing material into the public domain.

Conclusion

Intuitively, something seems amiss in the *Riley* case. Murray interpreted an ambiguous statement in a particular way and criticised it in strong terms. The judge found, as a matter of fact, that Riley's tweet referred to the very incident Murray interpreted it as referring to, but that it could have borne more than one meaning in relation to that incident. Murray thereby lost the case because she could not prove the truth of her apparently factual allegation, having been required to do so.

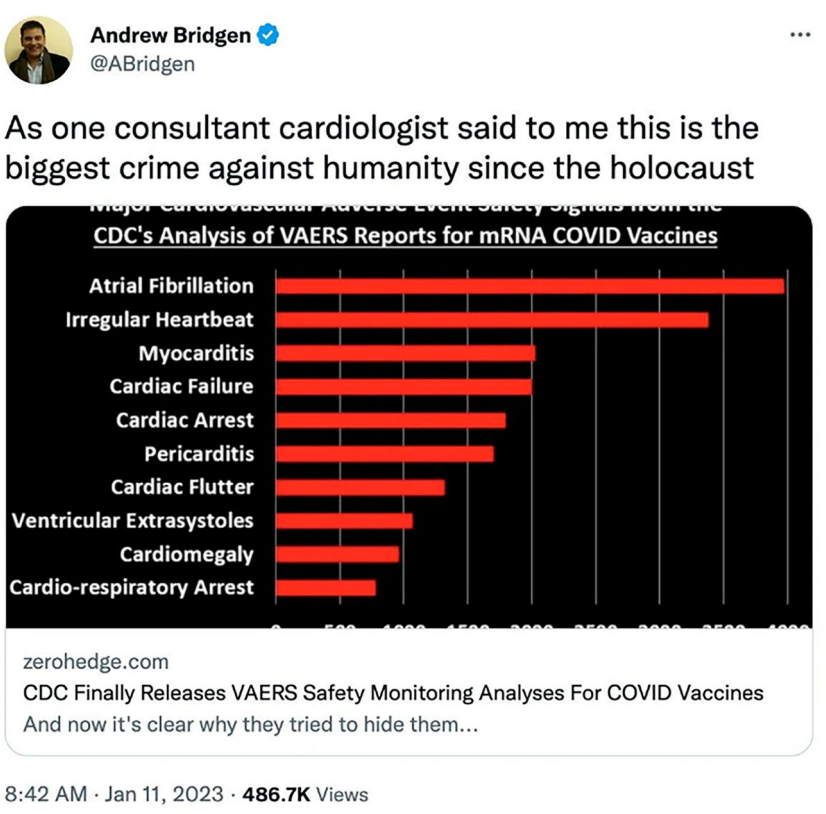
But all statements are ambiguous. Reader response theory tells us this. Some statements are perhaps more obviously ambiguous (in the minds of some interpretative communities) than others, but all statements are amenable to multiple interpretations. This is how language works. Law struggles to cope with this reality, for sure. This is because, as Stanley Fish puts it, 'the law wishes to have a formal existence'.[82] Judges, whose interpretative community

[82]Stanley Fish, 'The Law Wishes to Have a Formal Existence' in *There's No Such Thing As Free Speech* (OUP 1994).

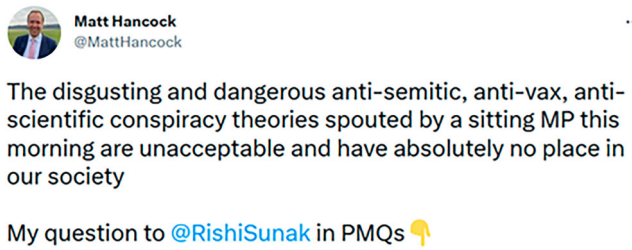
is embedded tightly within their experience of the highly formalised world of legal practice, unsurprisingly also struggle with this from time to time. But they could cope better if the proposed rule were to be adopted.

The only real fact in this whole affair is the fact that Murray's statement was a statement of opinion. Not just the bit at the end about Riley being dangerous and stupid, but the crucial part where Murray interprets Riley as having legitimised a violent attack on Corbyn. A proper understanding of reader response theory, including that theory's clear relevance to defamation law, leads to that conclusion. A proper reading of s.3 Defamation Act 2013 also leads to that conclusion. And, even if what I say is a proper reading of s.3 is wrong, then a contextual understanding and application of the old *Kemsley* rule, taking into account the particular characteristics of the ordinary, reasonable Twitter user, also leads to the same conclusion.

Riley, then, is deeply unsatisfactory. But what is most troubling is that it is not confined to itself. For the fact pattern that gave rise to *Riley* can and will be repeated in the future. It is being repeated, daily, across social media – particularly Twitter. For example, on 11 January 2023, the (now former) Conservative Party MP, Andrew Bridgen, posted this tweet, which linked to a webpage promoting anti-vaccination conspiracy theories:



Bridgen's parliamentary colleague, Matt Hancock MP, quickly criticised this tweet with one of his own, in which he accused 'a sitting MP this morning' of spouting 'antisemitic, anti-vax, anti-scientific conspiracy theories', and included a video clip of himself asking a question at Prime Minister's Questions (in which he refers only to such theories being 'promulgated online this morning' and does not identify Bridgen either by name or role).[83]



Note that nothing in Hancock's tweet references directly or links to Bridgen's tweet. Yet Bridgen commenced libel proceedings against Hancock, presumably on the basis that he (Bridgen) was identified by way of legal innuendo. Here, then, we see a political polemicist, Bridgen, making a statement that might be (and was by some) interpreted as anti-Semitic (in that it invoked the holocaust in a way that trivialises it), being called out for doing so, and commencing libel litigation in response. Based on the ruling in *Riley*, Hancock might well have to defend such a claim by pleading the defence of truth. Moreover, *anybody else* who interpreted Bridgen's tweet in the same way and posted their own criticism without quote-tweeting Bridgen or linking to his original tweet could quickly find themselves in the same position.[84]

[83] <https://twitter.com/MattHancock/status/1613159691785482240> (accessed 25/4/23).

[84] For another (less politically-charged) example, consider the recent furore involving the former world chess champion, Magnus Carlsen, and the (significantly lower rated) young American Grandmaster, Hans Niemann. In 2022, Carlsen unexpectedly lost to Niemann in a game at the Sinquefield Cup. Carlsen subsequently withdrew from the remainder of the tournament, but gave no direct explanation of his decision to do so. Instead, on his Twitter account, Carlsen tweeted an old clip of the well-known football coach, José Mourinho, speaking to camera after a football match. In this clip, Mourinho said that he preferred not to speak (about some controversial refereeing decisions) because 'if I speak I am in big trouble' (The video clip is no longer accessible via Carlsen's original tweet, due to a copyright complaint. But the same clip can be viewed here: https://youtu.be/9wtvXoXh0VU (accessed 25/4/23).) What, then, were Carlsen's 836,000 Twitter followers to make of this tweet? Numerous of his fans interpreted it as insinuating that Carlsen believed Niemann to have been cheating during their game. Some weeks later, Carlsen did in fact make allegations that more explicitly assert that something was amiss about Niemann's performance in that game (https://www.theguardian.com/sport/2022/sep/27/magnus-carlsen-hans-niemann-chess-cheating (accessed 25/4/23)), and Niemann has filed a defamation suit in the US against Carlsen in respect of these more explicit allegations (https://www.courthousenews.com/wp-content/uploads/2022/10/niemann-carlsen-chess-complaint-usdc-missouri.pdf (accessed 25/4/23)). However, it is entirely possible, owing to the approach taken in *Riley*, that Carlsen might instead have issued libel claims against anyone and everyone who interpreted his ambiguous Mourinho-clip tweet as meaning that he was accusing Niemann of cheating (as many of

Defamation law must not continue to deal with such cases in this way. For to do so will lead to chilling effects on public discourse – particularly on matters of political controversy and significance – and to significant logistical problems for the courts. Fortunately, there is a simple enough solution. The courts should simply recognise that a statement which expressly or impliedly interprets an earlier statement is itself a statement of opinion. The rest can be left to the elements of the defence of honest opinion, which defendants will either be able to make out or not.[85]

Acknowledgement

I am grateful to Jeevan Hariharan, Andrew Kenyon, Richard Mullender, Matthew Nicklin, and one anonymous reviewer for helpful comments upon, and criticisms of, earlier drafts.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Notes on contributor

Thomas DC Bennett is a senior lecturer in law at City, University of London. He has written extensively on the law relating to privacy, defamation, and judicial methodology in elaborating the common law. He is the cofounder and co-host of *The Media Law Podcast*.

Carlsen's own supporters did, vociferously piling in against Niemann). For, he could argue, it is defamatory of him to suggest he would publicly, and without providing evidence, accuse a fellow professional of cheating. And, based on *Riley*, anyone who did so without quote-tweeting the Mourinho-clip tweet might well be held liable in the English courts. That this did not happen owes nothing whatsoever to English defamation doctrine, and everything simply to Carlsen's apparent non-interest in issuing such claims.

[85]For the avoidance of doubt, absolutely everything I have said in this essay is just my opinion.