Reflective Equilibrium: Individual or Public?

Abstract: The paper explores whether the method of reflective equilibrium (RE) in ethics and political philosophy should be individual or public in character. I defend a modestly public conception of RE, in which public opinion is used specifically as a source of considered judgements about cases. Public opinion is superior to philosophical opinion in delivering judgements that are untainted by principled commitments. A case-based approach also mitigates the methodological problems that commonly confront efforts to integrate philosophy with the investigation of popular attitudes. This conception of RE is situated in relation to alternative accounts, including those of Daniels, Rawls and Wolff and de-Shalit.

Key words: Reflective equilibrium, Ethics, Political Philosophy, Public Opinion, Justification, Expertise

When they characterize their method, moral and political philosophers often appeal to the idea of reflective equilibrium (RE): a process of moving back and forth between principles and considered judgements, revising each where necessary in order to achieve the firmest possible system of mutual support between them.[1] This paper addresses one underexplored question about the nature of RE: namely whether this method should be individual or public in character.[2] Whose judgements and principles count as legitimate inputs into RE? And who is to be in charge of the process of revising and reconciling these elements? In particular, is RE a matter for the individual philosopher, who works between her own judgements and principles, seeking feedback where necessary from

[1] Norman Daniels, "Introduction: Reflective equilibrium in theory and practice," in his *Justice and Justification: Reflective Equilibrium in Theory and Practice* (Cambridge: Cambridge University Press, 1996), pp. 1-17, at p. 2. For Daniels, background theories also play a crucial role in his preferred approach of wide reflective equilibrium. See section 1 for further discussion of the distinction between narrow and wide RE.

[2] There are few extended discussions of this question in the literature on RE. Exceptions include Catherine Z. Elgin, *Considered Judgment* (Princeton, NJ: Princeton University Press, 1996), pp. 111-119 and Susanne Hahn, *Überlegungsgleichgewicht(e): Prüfung einer Rechtfertigungsmetapher* (Freiburg/München: Alber Thesen, 2000), pp. 58-59 and pp. 123-127.

academic colleagues? Or is it a practice that should, in some way, be opened up to a wider, non-philosophically trained, public.[3]

The paper outlines some alternative ways of understanding the ambiguous notion of public reflective equilibrium (PRE). I reject a strong conception of PRE on which the philosopher and the public each contribute both principles and judgements and they work in partnership to revise and reconcile these inputs. The paper argues that this approach is too ambitious about the contribution that public opinion can make to the practice of moral and political philosophy. However, I also resist a fully individualist approach, on which the philosopher has no use for wider public views. Rather, it is argued that RE should be public in a modest sense: public responses to philosophical cases represent an important input on the judgements side. Public opinion is superior to philosophical opinion in better answering the demand of RE for judgements about cases that are untainted by principled commitments. Such an approach also mitigates the methodological problems that commonly afflict efforts to integrate philosophy with the investigation of popular attitudes. This conception of RE is situated in relation to alternative accounts put forward by theorists including Daniels, Rawls and Wolff and de-Shalit. The question addressed here is conditional in character: *If* we think that justification in moral and political philosophy involves seeking coherence among judgements and principles, how, if at all, should we input public opinion? Thus I seek to show what follows for the role of public opinion if we are committed to the method of RE, rather than directly to defend that method against external critics.

The paper proceeds in five main parts. Section 1 addresses some background questions about the nature and purpose of RE and thereby clarifies the scope of my argument. Section 2 briefly outlines four existing accounts of RE. I characterize each of these accounts as either individual or public

[3] I do not mean to imply that philosophy, as traditionally practised, is a solipsistic activity; of course philosophers do routinely talk to colleagues about their work (although it is not clear that this typically amounts to engaging in RE together). In asking whether RE should be individual or public, I am concerned primarily with whether RE should be opened up to a non-philosophical audience. See section 4.A for further discussion.

across three distinct dimensions. Section 3 argues that there is value in "going public"[4] in RE, whilst rejecting the strong notion of PRE proposed by Wolff and de-Shalit in *Disadvantage*. Section 4 defends a specifically case-based understanding of the public component of RE. The arguments of section 4, I suggest, place the burden of proof on those who would restrict RE to the judgements of philosophers. Section 5 addresses the objection that reflective equilibrium demands *considered* judgements, and that such judgements are more likely to be reached by those with philosophical expertise. I argue that the countervailing case in favour of the greater reliability of philosophers' judgements is insufficiently strong to discharge the burden identified in section 4.

1. The Aims of Reflective Equilibrium

Before addressing the question of the role of public opinion in RE, it is essential to consider what RE is *for*; how we should approach RE will turn on what the method aims to do.[5] Although it has applications in many areas of philosophy, I focus here on RE as a method of justification in ethics and political philosophy.[6] But even within these domains, there are several important junctures at which we can differently conceive of the role or aims of RE. First, what is the status of the outputs of RE? Is RE a tool that seeks to track independent moral truths, or is it a constructivist procedure that generates those truths?[7] Or should we bracket the question of objective moral truth altogether for the purposes of engaging in RE?[8] Moreover, we might think that there is a distinctive role for RE in political philosophy, as a device of *political* constructivism. Understood in this way, the point of RE is

[4] I take this phrase from Elgin, *Considered Judgment*, p. 111.
[5] On the dependence of the approach on the aims of RE, see Hahn, *Überlegungsgleichgewicht(e)*, p. 126.
[6] The idea of RE is often traced to Goodman, in the context of inductive logic – see his *Fact, Fiction, and Forecast* (Cambridge, MA: Harvard University Press, 1955).
[7] For useful related discussion of the distinction between perfect, imperfect and pure procedures, see Elgin, *Considered Judgment*, Chapter 1. For further discussion of RE understood in constructivist terms, see François Schroeter, "Reflective Equilibrium and Antitheory," *Noûs* 38 (2004): 110-134.
[8] John Rawls, "The Independence of Moral Theory," *Proceedings and Addresses of the American Philosophical Association* 48 (1974-5): 5-22, at pp. 9-10.

not to track independent moral truths, but rather to produce an overlapping consensus that has authority by virtue of being such.[9]

For the purposes of the paper, I assume that RE aims at independent moral truths (although it may not succeed)[10] and I set aside the notion of RE as a device of political constructivism. In other words, I treat RE as an epistemic, rather than a democratic tool. Although I cannot develop my argument here, I do not believe that RE is plausible as an account of political constructivism. For example, whilst securing legitimate political decisions often depends on obscuring deeper divisions, the searching and iterative nature of RE exposes potential sources of disagreement. [11] Thus the remainder of the paper does not directly address the political module of RE in Rawls' later work. Rather I offer a defence of a moderately public approach to RE that is independent of the motivating concerns of *Political Liberalism*.[12]

Secondly, RE can be thought about both as an account of how we justify our moral and political theories, and as a description of the process by which we discover them.[13] The specifically case-based approach defended here concerns RE as a justificatory device. Thus my argument is compatible with the view that public opinion, in a wider variety of forms, will sometimes be an

[9] I thank an anonymous reviewer for this formulation of RE as political constructivism and for pressing me to say more about this issue. Understood in this way, the case for going public in RE turns on concerns of political legitimacy, rather than epistemic considerations. I criticize legitimacy-based arguments for an opinion-sensitive approach to political philosophy elsewhere – see Alice Baderin, "Political Theory and Public Opinion: Against Democratic Restraint," *Politics, Philosophy and Economics* (2015) DOI 10.1177/1470594X15621044

[10] This is widely (although not universally) shared view. See, for example, Norman Daniels, "Wide Reflective Equilibrium and Theory Acceptance in Ethics," *The Journal of Philosophy* 76 (1979): 256-282, at pp. 276-280; Michael DePaul, "Methodological Issues: Reflective Equilibrium," ed. Christian Miller, *Continuum Companion to Ethics*, xciii.

[11] On the possibility of agreement on practical issues despite deeper value disagreement, see Cass Sunstein, "Incompletely Theorized Agreements in Constitutional Law," *Social Research: An International Quarterly* 74 (2007): 1-24. For further critique of the political-constructivist understanding of RE, see Margaret Holmgren, "Wide Reflective Equilibrium and Objective Moral Truth," *Metaphilosophy* 18 (1987): 108-124, at pp. 116-124.

[12] See Section 2 for further discussion of Rawls' changing conception of RE. There is also some ambiguity in Wolff and de-Shalit's conception of RE: they appeal to epistemic considerations, but also to concerns of political feasibility and democracy in their defence of PRE. Jonathan Wolff and Avner de-Shalit, *Disadvantage* (Oxford: Oxford University Press, 2007) – compare p. 41, p. 98; p. 3, p. 21; p. 41. Here I focus exclusively on the epistemic strand of Wolff and de-Shalit's argument.

[13] Cf. John Rawls, "Outline of a Decision Procedure for Ethics," *The Philosophical Review* 60 (1951): 177-97, at p. 196. For a related distinction between "context of discovery" and "context of justification", see Herman Cappelen, *Philosophy Without Intuitions* (Oxford: Oxford University Press, 2012), p. 230.

important creative or imaginative resource for philosophy. Thirdly, RE is commonly understood both as a practicably workable method and as an unobtainable ideal of end-state justification. The paper assumes that RE should meet at least a weak realism constraint: "An account of methods can of course involve some idealizations, but it runs the risk of being pointless if it loses contact with our actual cognitive powers."[14] Thus I argue that strongly public RE is problematic in part because it fails to meet this modest feasibility condition.

Finally, a significant amount of debate around RE has centred on the distinction between wide and narrow conceptions. Daniels, in particular, has emphasized the move to wide RE as a response to the charge that RE is overly conservative. On this account, RE takes in background theories (moral and nonmoral), as well as a broad range of candidate principles. For example, Rawls' arguments about the original position are wide, because they appeal to background ideas about the nature of free and equal persons and the idea of a well-ordered society.[15] Some accounts of RE suggest that its materials are broader even than Daniels acknowledges. For example, DePaul emphasizes the role of personal experience, of art, literature and life, as a legitimate component in RE.[16] Wide RE looks appealing, to the extent that it increases the justificatory resources available to us. However, it is unclear how much work the move to wide RE can do in the face of very fundamental objections to RE as a coherentist form of justification.[17] Within political philosophy, wide RE also involves us in a complex debate about the proper place of empirical considerations in theorizing about justice. Although I do not discuss the role of background theories, the paper is neutral between narrow and wide accounts, because the reconciliation of principles and judgements retains an important role

[14] Schroeter, "Reflective Equilibrium and Antitheory," pp. 122-23.
[15] Daniels, "Wide Reflective Equilibrium," p. 260.
[16] Michael R. DePaul, *Balance and Refinement: Beyond Coherence Methods of Moral Inquiry* (Florence, KY: Routledge, 1993), Chapter 4.
[17] Much debate centres on the value of Daniels' "independence constraint": that the background theories should be based on an independent set of judgements – see his "Wide Reflective Equilibrium," p. 260. For discussion, see Carl Knight, "The Method of Reflective Equilibrium: Wide, Radical, Fallible, Plausible," *Philosophical Papers* 35 (2006): 205-229, at pp. 210-211 and Daniel Little, "Reflective Equilibrium and Justification," *The Southern Journal of Philosophy* 22 (1984): 373-387, at pp. 380-83. A further worry is that the move to wide RE defeats the conservatism charge at the cost of rendering RE vacuous – see T. M. Scanlon, "Rawls on Justification," in *The Cambridge Companion to Rawls*, ed. Samuel Freeman (Cambridge: Cambridge University Press, 2002), pp. 139-67, at p. 150-5.

within wide RE. Thus, depending on the stance we take on the wide versus narrow question, the modestly public approach defended here can be seen either as exhaustive of RE, or as an element (perhaps the first stage) within a broader process.

2. Mapping the Terrain

The table below summarizes four accounts of RE that embody different answers to the individual versus public question.

		Individual (Daniels, DePaul, Scanlon)	**Strongly public (Wolff and de-Shalit)**	**Moderately public: principled (Miller)**	**Moderately public: case-based (Early Rawls)**
Inputs	**Principles**	Philosopher	Philosopher & public	Public	Philosopher
	Considered judgements	Philosopher	Philosopher & public	Philosopher	Public
Process		Philosopher	Philosopher (& public?)	Philosopher	Philosopher

I distinguish three dimensions across which RE might be understood in individual or public terms: Who is the source of two key inputs and who is to carry out the process of revising and reconciling these elements? Here I understand the distinction between principles and judgements in terms of the relatively general versus the particular. Whilst this interpretation is commonplace in the literature on RE, it is sometimes argued that considered judgements can also be general in form. Notably, Rawls (after "Outline of a Decision Procedure") suggests that we can have judgements "at all levels of generality."[18] However, it is insufficiently clear how we are to understand the distinction between principles and judgements once we separate it from the particularity issue. Most commonly, the category of judgements is redescribed in terms of confidence: RE is about

[18] Rawls, "The Independence of Moral Theory," p. 8. See also T. M. Scanlon, "The Aims and Authority of Moral Theory," *Oxford Journal of Legal Studies* 12 (1992): 1-23, at p. 9; Georg Brun, "Reflective Equilibrium Without Intuitions," *Ethical Theory and Moral Practice* 17 (2014): 237-52, at pp. 239-40. I thank an anonymous reviewer for drawing my attention to this interpretation of the principles/judgements distinction.

systematizing *our most firmly held judgements*, at whatever level.[19] But this commits us to the problematic view that we should filter out, at the outset, all those judgements in which we are less confident. As Carl Knight persuasively argues, such a filter goes too far in pre-empting the process of reflective equilibrium.[20] Indeed, theorists who express the distinction between judgements and principles in terms of confidence sometimes also reject the confidence criterion for considered judgements, creating an unacknowledged tension within their account.[21] For this reason I prefer the general versus particular form of the distinction.

With this clarification in place, we can turn to the table. The idea of RE has been developed most extensively in ethics and political philosophy by Norman Daniels. According to Daniels, "The key idea underlying this view of justification is that we "test" various parts of our system of beliefs against the other beliefs we hold, looking for ways in which some of these beliefs support others, seeking coherence among the widest set of beliefs, and revising and refining them at all levels when challenges to some arise from others."[22] But from whom is this system of beliefs derived? And who is responsible for testing the different elements against each other? Although Daniels appeals to a "we" here, a closer reading suggests that his conception of RE is individualist. For example, he claims that RE aims "to produce coherence in an ordered triple of sets of beliefs *held by a particular person*."[23] Daniels also rejects the idea that we should admit only judgements on which there is widespread public agreement into the initial set of considered moral judgements. He concludes that such an approach seeks consensus too early in the process: it "may shift, in too crude a fashion (losing too many possibilities), the intermediate conclusions of my procedure into the position of methodologically warranted starting points."[24]

[19] For example, Scanlon, "Aims and Authority," p. 9.
[20] Knight, "The Method of Reflective Equilibrium," p. 208.
[21] For example, compare Schroeter, "Reflective Equilibrium and Antitheory," p. 117 with footnote 21, p. 131.
[22] Norman Daniels, "Reflective Equilibrium," in *Stanford Encyclopedia of Philosophy* (Winter 2013 edition), ed. Edward N. Zalta. Available at: <http://plato.stanford.edu/archives/win2013/entries/reflective-equilibrium/>.
[23] Daniels, "Wide Reflective Equilibrium," p. 258.
[24] Ibid, p. 282.

Thus on Daniels' account RE is an essentially personal process: it works with the views of, and is carried out by, the individual philosopher. This individualist account is the most widely adopted in the philosophical literature. For example, DePaul emphasizes that "Cases of conflict between a provisional theory and a considered moral judgement must be resolved individually; the inquirer must choose what to revise solely on the basis of what she is most strongly committed to, on the basis of what seems to her to be most likely to be true."[25] Similarly, for Scanlon, "the search for Reflective Equilibrium is essentially a first-person enterprise."[26]

A very different model of RE has been defended more recently by Jonathan Wolff and Avner de-Shalit. Wolff and de-Shalit carried out a number of in-depth interviews and used the findings to shape their theory of disadvantage. They term this empirically-informed approach to political philosophy "public reflective equilibrium":

> "instead of using the well known 'reflective equilibrium' technique alone – trying to balance a particular philosopher's theory and intuition – we brought … interviewees into the process. Accordingly we did more than simply learn about people's attitudes and views. Instead, we consulted our interviewees about our views and we learnt from them. In this process we revised and modified our theory according to the theories and intuitions expressed by the interviewees."[27]

A crucial feature of Wolff and de-Shalit's account is that it accords the philosopher and the public symmetrical roles in RE, in two respects. First, both input the same set of materials. Thus Wolff and de-Shalit are concerned to distinguish their approach from the idea of "contextual reflective equilibrium," which they detect in Walzer's work. In Walzer's model, they suggest, the intuitions are

[25] DePaul, *Balance and Refinement*, p.19.
[26] Scanlon, "Aims and Authority," p. 10.
[27] Wolff and de-Shalit, *Disadvantage*, p. 12. Wolff and de-Shalit refer here to "intuition," rather than judgements. I use the broader language of judgements throughout in order to sidestep the extensive debate about the precise definition of intuition. However, to the extent that intuition is typically understood as immediate or noninferential, there is an affinity between accounts of RE that require that judgements be intuitive and my defence of the principle-independence requirement in Section 4.

those of the philosopher and the "community's intellectuals," whilst the theories are those of the philosopher alone. In contrast, in PRE, "the theories considered are also those of the public, as well as the philosopher."[28] At times, Wolff and de-Shalit suggest that it is then up to the philosopher to fashion these inputs into a coherent whole. However, elsewhere they describe their respondents as active participants in the reasoning process. For example, they call for political philosophy "to start with the general public, activists and individuals *who are engaged in moral reasoning* and their dilemmas, alongside the theory and reasoning of the political philosopher."[29] On this latter account, the roles of the philosopher and the public are symmetrical in a second sense: the public do not simply contribute the raw materials for RE, they also engage as partners with the philosopher in the process of doing RE.[30]

Like Wolff and de-Shalit, David Miller sees an important role for public opinion within political philosophy. However, Miller argues that we should concentrate on uncovering the *general principles* that structure public attitudes: "the evidence [about public opinion] helps to confirm the theory *by revealing which principles people do in fact subscribe to*. The aim is to achieve an equilibrium whereby the theory of justice appears no longer as an external imposition conjured up by the philosopher, but as *a clearer and more systematic statement of the principles that people already hold*."[31] He suggests that the philosopher should be less concerned with popular responses to concrete cases; these are viewed as more superficial attitudes that can legitimately be set aside.[32]

[28] Ibid, p. 42.
[29] Ibid, p. 43. Emphasis added.
[30] Note that, although they characterize their approach as public reflective equilibrium, Wolff and de-Shalit's respondents were not a random sample of the general public. Rather, they selected participants with personal experience of disadvantage and professionals working with disadvantaged groups. This feature of their approach, which I set aside for the purposes of the paper, raises some interesting further possibilities. For example, as an anonymous reviewer has suggested to me, we might draw on the empirical knowledge of experiential experts to identify problematic practical consequences of candidate philosophical principles.
[31] David Miller, *Principles of Social Justice* (Cambridge, Mass.: Harvard University Press, 1999), p. 51. Miller, in one place, suggests something closer to the approach I advocate here, when he puts forward the "requirement that there should be a reasonably close fit between the theory and our [i.e. the general public's] pre-theoretical *considered judgements*" - see his "Two ways to think about justice," *Philosophy, Politics & Economics* 1 (2002): 5-28, at p. 6. However, Miller does not directly discuss the difference between drawing on popular principles and judgements, and his account predominantly emphasizes the former.
[32] For example, Miller, "Two ways to think about justice," p. 7.

Thus Miller offers us a third possible account of RE: one on which public attitudes are taken seriously on the principles, but not on the judgements side.

The final possibility outlined in the table is the opposite of Miller's approach: we use public opinion as a source of judgements, but not of principles. I have identified the early Rawls with this approach. However, Rawls cannot be fitted neatly into the table; depending on where we look, we can find more individualist or more public accounts of RE in his work. For example, the former is suggested in a *Theory of Justice* when Rawls tell us that "If we can characterize one (educated) person's sense of justice, we might have a good beginning toward a theory of justice. We may suppose that everyone has in himself the whole form of a moral conception. So for the purposes of this book, the views of the reader and the author are the only ones that count. The opinions of others are used only to clear our own heads."[33] A more public conception of RE is implied when Rawls characterizes the moral theorist as "an observer, so to speak, who seeks to set out the structure of other people's moral conceptions and attitudes."[34]

The picture is complicated further when we consider how Rawls' conception of RE changes with his growing sensitivity to the fact of reasonable pluralism and the politicization of his theory. Daniels emphasizes that wide RE as an individual process remains in place in *Political Liberalism*. Rawls adds a second layer of "political reflective equilibrium": a more restricted version of RE that addresses questions about constitutional essentials and matters of basic justice using only the materials of public reason.[35] When individuals who hold diverse comprehensive views each accept justice as fairness as an element within their various wide reflective equilibria, then we can be said to have achieved "general" RE.[36] Thus, by *Political Liberalism*, there are two levels or forms of RE: one

[33] John Rawls, *A Theory of Justice* (Cambridge, Mass.: Belknap Press of Harvard University Press, 1999), p. 44.
[34] Rawls, "The Independence of Moral Theory," p. 7. Although Rawls goes on to suggest that, in practice, the best way of studying people's views is to look to the main conceptions represented in moral philosophy.
[35] Norman Daniels, "Reflective equilibrium and justice as political," in his *Justice and Justification: Reflective Equilibrium in Theory and Practice* (Cambridge: Cambridge University Press, 1996), pp. 144-75, at p. 153.
[36] Ibid, p. 149. Political RE is public in the sense that it works with shared ideas from our political culture. It is less clear whether the *process* of political RE is a public or an individual one – cf. Daniels, "Reflective Equilibrium."

addressed to a restricted set of political questions and the other addressing questions of morality more broadly.

As discussed in the previous section, I do not engage directly with the specifically political component of RE in Rawls' later work. Instead the paper addresses the surviving individual module in Rawls' later work, to show that the later Rawls should, in one respect, go *more* public in RE (independently of the politicization of his theory). Specifically, I build on a suggestion in Rawls' earliest account of RE: that "the considered judgements of competent judges are the most likely repository of the working out of men's sense of right and wrong, a more likely one, for example, than that of any particular individual's alone."[37] Here, Rawls points to the possibility of a specifically case-based approach to public opinion. A version of this early Rawlsian conception of RE will be elaborated and defended in sections 4 and 5.

3. Neither Fully Individual Nor Fully Public

This section argues that there are good reasons to accord public opinion a role in RE, but there are also significant difficulties involved in going public and therefore limits to the role that popular attitudes can play. First, I suggest that Wolff and de-Shalit's conception of RE is too ambitious about the contribution that public opinion can make to moral and political philosophy. Secondly, I criticize the dominant individualist conception of RE.

A. Strongly Public Reflective Equilibrium is too Ambitious about the Contribution of Public Opinion

Wolff and de-Shalit's description of their method seems to give a fundamental role to public attitudes. However, further work is needed to unpack the deceptively simple idea of PRE and to clarify precisely how public and philosophical views are to be combined on this account. For example, do we first try to integrate the principles of both the public and the philosopher, before weighing them against the combination of their considered judgements? Or should we undertake a

[37] Rawls, "Outline of a Decision Procedure," p. 189. As noted earlier, Rawls subsequently says that judgements can also be general in form. But here he uses the term to refer specifically to judgements about cases.

process of RE within public opinion, before weighing popular views against the philosopher's own convictions? It is unclear exactly how PRE is supposed to work, with so many inputs in play. Without an answer to these questions, PRE stands for a broad intention to take public opinion more seriously in political philosophy, rather than a concrete method of justification.

As well as being problematically vague, this approach is too ambitious about the contribution that popular attitudes can make to philosophy. First, it demands that the public actually possess relevant principles. But even where the public do hold ideas that bear on philosophical debate, this typically does not involve anything as structured as a moral or political principle. This is evident, for example, in a study that set out to explore public perceptions of citizenship in relation to liberal, cultural pluralist and communitarian models. The researchers report that although people can understand the notion of citizenship when it is put to them, it has less force than expected as an organizing framework for everyday political thought.[38] Similarly, participants in Wolff and de-Shalit's study had many ideas to offer about the nature of disadvantage. However, they often expressed these views in ways that blurred the boundary between resources and capabilities and between capabilities and functionings, suggesting that these key theoretical distinctions held little salience for them.[39]

Philosophers sometimes write as if society at large is populated by adherents of various ethical world views: "Some … [citizens] believe that just policies are those which maximize overall utility within the constraints of fundamental rights; others that they are those which maximize average utility; others still think that laws are just in virtue of their conformity with Kant's Categorical Imperative or Rawls's principles of justice; some religious citizens hold that justice requires laws and policies to reflect our status as God's creatures."[40] Perhaps this is just an unfortunate turn of phrase; if they were to reflect further, surely no philosopher really thinks that popular thought is consistently structured by Rawlsian, Kantian and various forms of utilitarian theory. However, it is

[38] Pamela Conover, Donald Searing and Ivor Crewe, "The Elusive Ideal of Equal Citizenship," *The Journal of Politics* 66 (2004): 1036-68, at pp. 1051-52.
[39] For example, Wolff and de-Shalit, *Disadvantage*, p. 57.
[40] Laura Valentini, "Justice, Disagreement and Democracy," *British Journal of Political Science* 43 (2013): 179-99, at p. 190. See also DePaul, "Methodological Issues," lxxvii.

important to expose and to challenge this picture, because it can operate implicitly to shape how we think about the relationship between philosophy and public opinion. Of course citizens do disagree about moral and political values, but simply imagining lines of theoretical contestation to be replicated among the wider population gives a misleading picture of public opinion and the ways in which it can inform philosophy.[41]

Secondly, strong PRE is problematic in light of evidence that people often hold divergent views at different levels. For example, one recent study reveals the complex structure of public attitudes towards desert. Freiman and Nichols find that people reject the *principle* that individuals can deserve extra rewards on the basis of lucky talents. However, when they are presented with *concrete cases* in which individuals perform differently solely because they have different natural talents, people tend to say that the more talented *do* deserve extra rewards.[42] Thus it appears that the public both believe and do not believe that individuals deserve rewards for lucky talents, depending on whether this is framed as an abstract or a concrete question. This kind of abstract/concrete effect within public opinion has been reported in a number of other areas, including attitudes towards immigration and moral responsibility.[43] From the point of view of RE, some inconsistency is not in itself a problem; after all there would be little for the method to do if all our views were already in harmony with each other. However, the messiness of public opinion does present the philosopher with the daunting task of working out which aspects of popular thought to revise or to give up. Wolff and de-Shalit do not offer an account of how we are to weigh and adjudicate between public opinion at different levels.

[41] Note that the issue raised here counts against a model (suggested to me by an anonymous reviewer) on which we do input public opinion on the principles side, but we give public principles lesser weight than philosophical principles. This discounting approach does not adequately address the problem that the public simply lack principled views on many moral and political questions.

[42] Christopher Freiman and Shaun Nichols, "Is Desert in the Details?" *Philosophy and Phenomenological Research* 82 (2011): 121-33.

[43] Toril Aalberg, Shanto Iyengar and Solomon Messing, "Who is a 'deserving' immigrant? An experimental study of Norwegian attitudes," *Scandinavian Political Studies* 35 (2012): 97-116; Shaun Nichols and Joshua Knobe, "Moral Responsibility and Determinism: The Cognitive Science of Folk Intuitions," *Noûs* 41 (2007): 663-85. For more general discussion of the abstract/concrete effect, see Walter Sinnott-Armstrong, "Abstract + Concrete = Paradox," in *Experimental Philosophy*, ed. Joshua Knobe and Shaun Nichols (New York: Oxford University Press, 2008), pp. 209-30.

A third difficulty arises on the view that the public also have a central role to play in RE as process. Later I will challenge the claim that philosophers possess special expertise in the formation of judgements about concrete cases. However, I do not wish to rule out all forms of philosophical expertise. In particular, we should expect philosophers to perform better than the general public when it comes to seeing the relationship between principles and cases. Such an understanding is essential to the process of revising and reconciling the different elements in RE. But to demand this ability of the general public is, to a significant extent, to ask that they *become* philosophers. Moreover, there are significant difficulties in conceiving of the intricate process of RE as a process of collective reasoning. How, as a collective matter, are we to determine what we ought to believe in the face of multiply conflicting judgements and principles?

PRE, in the strong sense defended by Wolff and de-Shalit, is both indeterminate and too demanding. The objection here is not simply that PRE is very challenging, or that its ideal end-state is not attainable. We should expect this of a method for justifying our beliefs about complex moral and political questions. The worry is that, as Wolff and de-Shalit describe it, RE is not a workable method. It is telling then that their actual research practice departs significantly from the way in which Wolff and de-Shalit characterize their approach. They claim that "we were open to be influenced by the interviewees, we were not assuming particular conceptions and understandings of these concepts [of functionings]; rather we wanted to reach full understanding on both sides – interviewer and interviewee."[44] This description of their approach suggests significant openness to the research encounter and a willingness to rethink their ideas in response to alternative views from participants. However, in practice, the important theoretical commitments are made *before* Wolff and de-Shalit engage with public opinion, and they are unwilling to disrupt their framework in significant ways. One category is added to their provisional list of capabilities in response to participants' views.[45] But more typically Wolff and de-Shalit reinterpret participants' responses in order to fit the empirical

[44] Wolff and de-Shalit, *Disadvantage*, p. 44.
[45] Specifically interviewees' emphasis on the importance of being able to communicate in the local language – ibid, pp. 60-61.

data into their pre-existing theoretical categories.[46] In this way, the symmetrical and partnership model of RE described in *Disadvantage* is not in fact realized. If the argument presented here is correct, this is not a contingent problem with their study; rather there are deeper reasons for Wolff and de-Shalit's departure from their professed method.

B. Against Individual RE

In light of the difficulties with strong PRE, should we retreat to an individual conception of RE? Why go public in RE at all? The answer to this question will emerge in part through my defence of a specifically case-based approach to public opinion in Section 4. However, it is useful first to identify some key arguments for going public in more general terms. As Wolff and de-Shalit argue, philosophers are a distinctive group. Thus evidence about wider popular attitudes can play an important role in identifying and correcting biases in our perspectives.[47] Comparing our views with those of the wider public allows us to identify those among our commitments that are widely shared. Where we find divergence, this forces us to consider why we hold the beliefs that we do and to reflect on how our views might have been shaped by our particular interests or social positions. Where we find public agreement, this also gives us valuable information: "people's commitments and experiences diverge. So a shared conviction is apt to have withstood a wider range of challenges than an idiosyncratic one."[48] In Elgin's terms, shared convictions have demonstrated greater "staying power"[49] and thereby enjoy enhanced epistemic status.

The general case for attending to wider public attitudes can be stated in uncontroversial terms. If we are dealing with complex moral issues, where our own experiences are particular or limited and self-interest can intrude, surely there is value in considering the views of others; particularly those who

[46] See their treatment of participants' ideas about the importance of "home" and views on education – ibid, p. 105 and p. 106.
[47] Wolff and De-Shalit, *Disadvantage*, p. 12.
[48] Elgin, *Considered Judgment,* p. 112. Whilst Elgin has a different view of the epistemic goals of RE to the one adopted in this paper (for Elgin, RE aims at understanding rather than truth and knowledge), her claims about the significance of agreement remain relevant.
[49] Ibid, p. 112. Elgin goes on to express reservations about the extent to which agreement, in general, adds to the tenability of our initial commitments. See sections 4 and 5 for further discussion of the (limits to the) significance of agreement.

are differently situated to ourselves.[50] But perhaps we can enjoy these epistemic benefits without going public in RE? Specifically, it is sometimes suggested that RE itself should remain individual (in terms of both inputs and process), but that we might subsequently compare our various individual equilibria to look for convergence.[51] Although the idea of individual RE followed by social comparison is commonly put forward by advocates of the individual approach, it has not been adequately developed. If a stable individual equilibrium is something we will never fully achieve (as defenders of individual RE acknowledge), it is unclear at what point we are to take the comparative step. Perhaps we should first do as much work as we can individually, before considering others' results. But what is the purpose of making the comparison? The most obvious answer is that, when we find disagreement this gives us reason to revisit, and possibly reconsider, our own view. But then we will want to try to pinpoint the sources of the disagreement. This will lead us to start to decompose the individual equilibria, to compare the different elements across equilibria to see where and why they diverge. In effect then we are still doing RE in the comparative stage, drawing on others' views as well as our own. In this way, the distinction between individual and public RE starts to break down. Once we admit the usefulness of comparing individual equilibria (for the reasons identified above), then RE itself takes on a public component.

4. For a Case-based Approach

There is value in going public in RE, but there are also problems with the strongly public approach advocated by Wolff and de-Shalit. Section 4 proposes a way forward, arguing that we ought to retain a narrower role for public opinion: as a source of judgements about cases. The argument proceeds in two parts. First, I suggest that public opinion better fulfils an important requirement of principle-independence. Secondly, I argue that engaging with public opinion in a case-based way mitigates the methodological problems that commonly arise when we seek to integrate philosophical inquiry with

[50] Alison Hills, "Moral Testimony and Moral Epistemology," *Ethics* 120 (2009): 94-127, at p. 123.
[51] For example, Daniels, "Wide Reflective Equilibrium," p. 281; Scanlon, "Rawls on Justification," p. 149.

the empirical investigation of popular attitudes. Thus, unlike a strongly public model, this moderately public approach represents a sufficiently realistic and workable model of RE.

A. Principle-Independence

The value of public judgements follows from a crucial observation about the requirements of RE: on this method, principles and judgements are meant, in the first instance, to represent *independent* inputs into the process of justification. We cannot use our case-specific judgements to justify our principles if we have already derived those judgements from the principles.[52] Thus from the perspective of RE, the philosophical naivety of public judgements is an asset. In contrast, it is much more difficult for the philosopher genuinely to separate her case specific judgements from her principled leanings. The claim here is not that the public do, or should, make judgements about philosophical cases without having in mind any reasons for those judgements.[53] But even if the public have some reasons in mind, they remain in a different (and, for the purposes of RE, advantageous) position relative to the philosopher. The public might have "proto-theoretical hunches about [for example] the reasons why it is morally wrong to break promises,"[54] but they are not aware of the key contenders in philosophical debate about the wrongness of promise-breaking, nor are they committed to some systematic philosophical view.

The priority of public responses rests on the idea that RE involves a requirement of principle-independence for judgements. What does this independence requirement amount to? Rawls' requirement that we exclude the *conscious* use of principles to determine judgements is necessary, but insufficient. If we have internalized a principle, we might recognize its implications in a particular case, without being aware that we are applying the principle. A judgement reached in this way fails

[52] Rawls, "Outline of a Decision Procedure," p. 183.
[53] Thus my argument does not depend on the conceptual possibility of brute moral judgement. For helpful discussion of this issue in light of developments in empirical moral psychology, see Jeanette Kennett and Cordelia Fine, "Will the Real Moral Judgment Please Stand Up: The Implications of Social Intuitionist Models of Cognition for Meta-ethics and Moral Psychology," *Ethical Theory and Moral Practice* 12 (2009): 77-96.
[54] Schroeter, "Reflective Equilibrium and Antitheory," p. 119.

to provide independent support for the principle.[55] Not all case-based judgements must fulfil principle-independence: It is essential to RE that we can revisit and revise judgements explicitly in light of a principle we come to find credible. However, the independence requirement does extend beyond only initial judgements. For example, during the process of RE, a principle is suggested which has implications not considered so far; new cases must therefore be judged. In this context, the independence requirement does apply and we should seek public opinion on the new cases.[56]

The problem of principle-bias is acknowledged in some recent literature on the role of intuitions, but it has not been adequately accommodated within accounts of RE.[57] For example, Raz frequently refers to the role of "pre-philosophical" judgements in RE, but he does not consider how *philosophers* are to generate these judgements. More generally, accounts of RE frequently describe a sequence in which we *first* identify a set of moral judgements, and then subsequently look for general principles that explain those judgements.[58] This suggests that our candidate principles come onto the table only after the relevant judgements are formed. But, on the assumption that RE is carried out by, and using the judgements of, an individual philosopher, these descriptions are misleading; the philosopher typically constructs and judges cases, or reflects on established cases, in light of her understanding of debates at the level of principle. These sequential accounts are telling even if they are not to be taken literally as an account of the order of justification, because they embody a principle-independence condition that their proponents implicitly rely on to make RE appealing.

---

[55] See Christopher Hitchcock, "Thought experiments, real experiments, and the expertise objection," *European Journal for Philosophy of Science* 2 (2012): 205-18, at p. 207.

[56] I thank an anonymous reviewer for pressing me to clarify my understanding of the independence requirement here.

[57] John Doris and Stephen Stich, "Moral Psychology: Empirical Approaches," in *Stanford Encyclopedia of Philosophy* (Winter 2012 edition), ed. Edward N. Zalta. Available at: <http://plato.stanford.edu/archives/win2012/entries/moral-psych-emp/>.; Hitchcock, "Thought experiments," p. 207 and pp. 212-3; Jeanine Weekes Schroer and Robert Schroer, "Two Potential Problems with Philosophical Intuitions: Muddled Intuitions and Biased Intuitions," *Philosophia* 41 (2013): 1263-81, at pp. 1274-79. For a rejection of the principle-bias concern, see Steve Clarke, "Intuitions as Evidence, Philosophical Expertise and the Developmental Challenge," *Philosophical Papers*, 42 (2013): 175-207, at pp. 184-6.

[58] For example, Scanlon, "Aims and Authority," p. 2; Schroeter, "Reflective Equilibrium and Antitheory," p. 116.

It is useful to unpack the problem of principle-bias further, since there are some distinct reasons why philosophers' judgements might be susceptible to (from the perspective of RE) illicit shaping by principles. First, philosophers will simply have greater awareness of how the cases relate to various candidate principles. As noted earlier, this awareness is essential when carrying out RE as process. However, it becomes problematic when it comes to generating the judgements themselves. Secondly, philosophers are more likely to be emotionally or professionally invested in particular principles.[59] There is strong evidence that the general principles people endorse are sometimes shaped by the concrete judgements to which they are already committed. This phenomenon has been termed "motivated moral reasoning": a process in which "an individual has an affective stake in perceiving a given act or person as either moral or immoral, and this preference alters reasoning processes in a way that adjusts moral assessments in line with the desired conclusion."[60] For example, individuals will recruit either consequentialist or deontological principles depending on which perspective supports their preferred moral judgements.[61] But philosophers are different from much of the general public in that they are sometimes emotionally invested in abstract theories or principles. Thus there is a danger of a reverse process of "motivated moral judgement" for the professional philosopher, in which her concrete judgements bend to her principled commitments. For example, Huebner and Hauser show how Thomson's treatment of concrete cases designed to elucidate the problem of altruistic self-sacrifice is governed, in a problematic way, by her principled commitments on this issue.[62]

Thirdly, we should consider whether there is anything unusual about the psychological make-up of philosophers, which might give rise to a distinctive tendency to approach concrete cases in a principled manner. Although there are, to my knowledge, no direct studies of this question, there is

[59] For a similar distinction between theory knowledge and "vested interest," see Hitchcock, "Thought Experiments," p. 213.
[60] Peter Ditto, David Pizarro and David Tannenbaum, "Motivated Moral Reasoning," *Psychology of Learning and Motivation* 50 (2009): 307-38, at p. 312.
[61] Ibid, p. 329.
[62] Bryce Huebner and Marc D. Hauser, "Moral judgements about altruistic self-sacrifice: When philosophical and folk intuitions clash," *Philosophical Psychology* 24 (2011): 73-94, at pp. 89-91.

some closely related evidence of a distinctive "philosophical temperament." Livengood et al. show that philosophers are typically more reflective: "they are less likely than their [equally well educated] peers to embrace what seems obvious without questioning it, and they are disposed to submit to scrutiny their intuitive inclination to judge that something is the case."[63] Whilst crucial to the effective conduct of many aspects of philosophy, these tendencies are again likely to militate against the production of principle-unmediated judgements about concrete cases. Each of these factors, particularly if they work in combination, makes it more likely that the philosopher, consciously or unconsciously, will filter her responses to cases through the lens of candidate principles.

Breaking down the problem of principle-bias in this way serves to problematize the category of "professional philosopher," which is typically left rather vague in discussions of philosophical expertise. For example, when it comes to political philosophy is the professional any career philosopher or only those with a specialism in political philosophy? Or, more narrowly, those who are research active within a particular area of debate, such as global justice or human rights? The extent to which the first and second mechanisms (awareness and emotional or professional investment in principles) apply depends on how broadly or narrowly we draw the parameters of this group. This suggests a possible revision to my argument. Perhaps rather than making RE (partly) public, we need only draw on the judgements of some subset of the philosophical community who do not work directly on the issues at stake?

Although this approach is likely to ameliorate the problem of principle-bias somewhat, best practice is to utilize the views of the general public. Christopher Hitchcock draws a useful analogy here with the way in which scientific experiments are conducted: independent coders are used to categorize observations, in order that coding is guaranteed not to be affected by knowledge of the aims of the experiment or by any investment in achieving particular results.[64] The thought here is that the

[63] Jonathan Livengood, Justine Sytsma, Adam Feltz, Richard Scheines and Edouard Machery, "Philosophical temperament," *Philosophical Psychology* 23 (2010): 313-30, at p. 314.
[64] Hitchcock, "Thought Experiments," pp. 212-3.

general public are in the best position to occupy the equivalent of the role of independent coders when it comes to philosophical thought experiments.

This section has argued that public opinion better answers the demand of the method of RE for principle-independent judgements about cases. It is important to emphasize the limited nature of the claim being made for public opinion, in two respects. First, the fact that public judgements fulfil the principle-independence requirement does not yet show that they have the necessary epistemic standing for inputs into RE. I address this issue in Section 5, through the lens of the considered judgements condition. Secondly, even setting aside the considered judgements requirement, popular judgements are not sovereign or decisive. As Rawls emphasizes, no element is beyond revision in an ongoing process of RE.[65] Thus the suggestion is that public opinion should take precedence over philosophical opinion when it comes to case-specific judgements; it is not that the philosopher should necessarily revise her theory in order to accommodate all such judgements. However, if she *consistently* refuses to accommodate her theory to public judgements, this starts to reveal something more general about the philosopher's approach: on the account presented here, it implies that she is not really serious about the method of RE itself. In other words, thinking about public opinion in the way suggested here can play a valuable role in forcing philosophers to consider how deeply they are committed to their professed mode of justification, with its central role for judgements about concrete cases. Thus the claim that public case-specific judgements enjoy priority by virtue of their relative philosophical naivety points to one of two possible conclusions: either we must sometimes revise our principled positions to accommodate contrary case-specific public judgements; or the method of RE, with its justificatory role for case-specific judgements, is not borne out in philosophical practice.[66]

B. Mitigating Methodological Challenges in the Integration of Philosophy and Public Opinion

[65] Rawls, "The Independence of Moral Theory," p. 8.

[66] For the related argument that philosophers do not, in fact, rely on intuitions despite frequently claiming to do so, see Cappelen, *Philosophy Without Intuitions*. For critical discussion of Cappelen, see Jennifer Nado, "The intuition deniers," *Philosophical Studies* 173 (2016): 781-800.

Challenges arise when we endeavour to bring public opinion to bear on ideas and principles from philosophy, whose purchase on everyday thought and discourse is often limited or uncertain. The wider public do hold views that are relevant to many areas of philosophical enquiry. However, they are not participants in those debates; they do not share the same language and, more fundamentally, they do not organize their ideas in the same way. In particular, philosophers often want to draw conceptual distinctions where there are no clear or stable lines in popular thought. For example, we commonly distinguish between the values of justice, humanity and efficiency and between normative and feasibility considerations. However, most people do not divide up their moral and political views like this; at least not in any consistent way. Philosophical debates are also typically conducted in terms much more abstract than those in which people formulate their beliefs. Thus philosophers who are interested in engaging with public opinion face some difficult issues about how they convey unfamiliar theoretical ideas in a research setting and ensure that participants offer their responses in terms that respect crucial distinctions.[67]

The case-based method of engaging with public opinion defended here serves to mitigate some of these difficulties, which were illustrated in relation to Wolff and de-Shalit's strongly public model. First, people generally find it easier to engage with unfamiliar philosophical ideas in the form of concrete cases, rather than abstract principles. For example, there is a growing body of work in experimental philosophy testing public views about various theories of knowledge. Epistemology appears to be just the kind of abstract, difficult and specialized subject with which outsiders would struggle to engage. Yet when experimental philosophers present non-specialist audiences with concrete cases embodying alternative theories of knowledge, respondents seem perfectly able to understand and to form judgements about them.[68] The more general reason for this is that the *cases*

[67] For helpful discussion of the difficulties involved in measuring popular views about justice, see Adam Swift "Public Opinion and Political Philosophy: The Relation between Social-Scientific and Philosophical Analyses of Distributive Justice," *Ethical Theory and Moral Practice* 2 (1999): 337-363, at pp. 342-44 and Miller, *Principles of Social Justice*, p. 47.

[68] See section 5 for discussion of some potential caveats. The example of epistemology suggests that public judgements need not be restricted to familiar, real or realistic cases – cf. Rawls, "Outline of a Decision Procedure," p. 5.

themselves might be unfamiliar, but the process of forming moral judgements about particular cases is a relatively familiar one. In contrast, abstract principles are doubly demanding: they are both unfamiliar in themselves and they demand that people *think in a way* that is out of the ordinary. We can draw an analogy here with the way in which linguists study popular views about grammar. They do not expect non-linguists to be able to engage with questions about grammatical principles, but they often do take seriously popular judgements about the grammaticality of concrete sentences.[69]

Secondly, on the modestly public account defended here, the process of RE is still owned and controlled by the individual philosopher; the public do not actually engage in RE, rather they provide material that assists the philosopher in undertaking that process. To that extent, RE remains an intra-personal process. Thus, I share with defenders of individual RE the view that the intricate process of RE is plausible only as an account of individual moral reasoning. [70]

Thirdly, the approach to RE defended here gives public opinion a significant role only on one side of RE. We use public judgements in a piecemeal manner as selected inputs into RE, thus we are not committed to working up popular thought into a coherent whole. It was noted in section 3 that popular attitudes sometimes diverge at abstract and concrete levels and this creates challenges for the strongly public conception of RE. If we take public opinion just as an input into one side of the process of RE, the tension between different elements of popular thought can be viewed more positively. The fact that, for example, the public tend to favour desert in concrete cases, whilst rejecting desert as a matter of principle, suggests that they are, at least, responding in a direct way to the cases; their case-specific judgements are not simply determined by prior principled commitments.

However, the problem of disagreement arises in a different form on the conception of RE defended here. Specifically, how do we deal with the fact that the public are likely to disagree in their judgements about particular cases? RE is not normally understood as a method for resolving

[69] Freiman and Nichols, "Is Desert in the Details?" p. 131.
[70] For example, DePaul, "Methodological Issues."

conflicts between judgements about the same case, so how are we to proceed?[71] The general case in favour of public over individual RE sketched in 3.B gives us reason to take the majority judgement as our starting point: more widely shared views are less likely to be the product of particular narrow or biased perspectives. However, this does not permanently bind us to the majority view, given that all elements within RE are open to revision. Indeed, the extent of agreement on a particular judgement is information that can be used during the revision process. Specifically, where disagreement is more extensive, we should give lesser weight to a particular judgement when it conflicts with a principle that we find plausible (although not discount the judgement entirely). There is significant indeterminacy in most accounts of RE when it comes to the question of what we should do in the face of conflicts between principles and judgements. [72] This is a recurring objection to RE, which I suggested occurs in particularly powerful form in relation to Wolff and de-Shalit's approach. The account I put forward does not fully resolve this indeterminacy, but it does add to the information we can draw on in the process of reconciling judgements and principles.

The approach to disagreement outlined here contrasts with a more stringent view, on which in the face of conflicting judgements that we are unable to explain away, we must discard those judgements altogether.[73] Such a response represents an overreaction to the phenomenon of disagreement on philosophical cases. As Williamson argues, "Disagreement can provide a reason to be somewhat more cautious then we might otherwise have been, in our handling both of eye-witness reports and of thought experiments; such caution is commonplace in philosophy. There is no need to be panicked into more extreme reactions."[74]

5. Objection: The Need for Considered Judgements

I have made two main claims in support of a moderately public conception of RE: first, public judgements best meet the requirement of principle-independence; secondly, a case-based approach

[71] I thank an anonymous reviewer for raising this question.
[72] DePaul, "Methodological Issues," lxxx. For DePaul's own response to this problem, see lxxxi.
[73] Cummins, "Reflection on Reflective Equilibrium," p. 115. For Cummins, this represents a fundamental objection to the method of RE.
[74] Timothy Williamson, *The Philosophy of Philosophy* (Malden, MA: Blackwell Publishing, 2007), p. 192.

mitigates the methodological challenges that arise when we seek to engage public opinion on unfamiliar philosophical terrain. These arguments place the burden of proof with those who would restrict RE to the philosopher's own judgements. This section considers, and rejects, an attempt to discharge this burden by appeal to philosophical expertise.

RE is typically understood to be concerned not with all and any judgements about cases, but rather with a subset of appropriately robust or reliable "considered" judgements. I have suggested that a non-philosophically trained public can deal reasonably well with philosophical thought experiments. But even if that is right, we might think that philosophers are *more* likely to have considered judgments. After all, philosophers spend a lot more time engaged in the activity of entertaining thought experiments; surely they will have some special expertise?[75] Perhaps then we should retreat from even the modestly public conception of RE defended here: "Although the philosophically innocent may be free of various forms of theoretical bias … that is not enough to confer special authority on innocent judgment, given its characteristic sloppiness."[76]

In order to address this objection, it is important to remind ourselves that a "considered judgement" cannot simply be whichever response coheres with our preferred principle: that would eliminate any independent role for case-specific judgements within the process of RE.[77] On Rawls' widely cited account, considered judgements require that: we are able to concentrate on the case at hand, for example we should not be in a highly emotional state; we do not stand personally to lose or gain; we are aware of the relevant facts. In other words, we pick out judgements made under conditions that

[75] For an appeal to philosophical expertise, see Timothy Williamson, "Philosophical Expertise and the Burden of Proof," *Metaphilosophy* 42 (2011): 215-29. For more extended (but critical) discussion of what philosophical expertise might consist in, see Jonathan M. Weinberg, Chad Gonnerman, Cameron Buckner and Joshua Alexander, "Are Philosophers Expert Intuiters?" *Philosophical Psychology* 23 (2010): 332-355. An alternative objection to my argument would concede that public judgements are superior to philosophers' judgements from the perspective of RE, but deny that we need to engage in empirical research to predict popular responses. Empirical evidence concerning philosophers' accuracy in forecasting public judgements is both limited and mixed. See Billy Dunaway, Anna Edmonds and David Manley, "The Folk Probably do Think What you Think They Think," *Australasian Journal of Philosophy* 91 (2013): 421-41; Justin Sytsma and Edouard Machery, "Two conceptions of subjective experience," *Philosophical Studies* 151 (2010): 299-327, p. 308.

[76] Williamson, *The Philosophy of Philosophy*, p. 191

[77] Rawls, "Outline of a Decision Procedure," p. 180.

avoid the most common sources of error.[78] On an individual conception of RE, I must ensure that these conditions are met in an introspective way, by inspecting my own circumstances and state of mind when I form my judgement. A public approach opens up possibilities for satisfying these criteria in a research setting: either by controlling the conditions of the research encounter itself, or by including filter questions to remove participants who fall short. As suggested in Section 2.B, evidence about patterns in public judgements can also help us to isolate distorting factors and thereby to pick out more reliable judgements. The claim here is not that some threshold of public agreement is necessary for a judgement to count as considered.[79] Rather, certain patterns in public opinion might give us reason to suspect that the conditions for considered judgement have not been met. For example, if we find that judgements are strongly linked to social class, we should consider whether self-interest is intruding.[80]

Thus we can apply some widely shared Rawlsian criteria for considered judgements to public judgements. Indeed given the ability to control the conditions of judgement in a research setting and the opportunity to learn from patterns in opinion data, a public approach has advantages over an individual introspective approach when it comes to the task of picking out considered judgements. However, a deeper worry about public judgements remains. We might be able to use empirical methods to discriminate among public judgements, but what about the fact that the philosopher will simply have thought more extensively and more carefully about the cases? Surely such reflection will result both in better understanding of the cases, and in a reduction of the influence of prejudice or social stereotypes on philosophers' responses?

Some recent evidence from experimental philosophy bears on the common-sense assumption that greater reflection leads to more reliable judgements. Schwitzgebel and Cushman compare order and framing effects in responses to a range of common thought experiments in ethics across two groups.

[78] Ibid, pp. 181-83 and *A Theory of Justice*, p. 42. Rawls also requires that considered judgements are held stably and confidently. As noted earlier (see footnote 19), we should resist these further filters on initial judgements.
[79] Cf. Rawls, "Outline of a Decision Procedure," pp. 182-183.
[80] Miller, *Principles of Social Justice*, p. 55.

Half of their participants were explicitly instructed to reflect before making their judgements, to consider potential arguments on both sides and to think about how they might respond to different variants of the scenario, as well as given a forced delay in responding.[81] Surprisingly, this "reflection condition" did not reduce the impact of apparently irrelevant features of framing and case-order on participants' responses. Research into the phenomenon of implicit bias also bears on the claim that, by thinking harder about the cases, philosophers will thereby reduce the effect of stereotypes or prejudices on their judgements. For example, Poehlman et al. report that the effect of implicit bias on behaviour is not reduced when that behaviour is more deliberate or controllable.[82]

In order for our judgements to count as considered, we need to have thought sufficiently to have grasped the case and to have taken it seriously. However, this does not mean that more reflection is better, in the sense that the more we reflect, the more considered are our judgements. Rather than simply insisting that philosophers' judgements are more reflective and therefore must be more considered, we would do better to try to identify some *specific* respects in which public judgements are more likely to be defective. Here I consider two potential tendencies for public misunderstanding of philosophical cases.

First, I wish to address an objection to relying on the judgements of non-philosophers pressed in a recent paper by Schroer and Schroer.[83] They argue that many such judgements are "muddled": that is, they are misdirected because they are not grounded in a proper understanding of the relevant conceptual landscape. For example, say we are looking for a judgment about moral responsibility. The non-philosopher may conflate this issue with the related, but distinct, question of justified punishment and thereby offer a judgement that misses the target. In contrast, the trained philosopher, "possesses a clear view of the complicated nexus of interconnected concepts of [the

[81] Eric Schwitzgebel and Fiery Cushman, "Philosophers' biased judgments persist despite training and reflection," *Cognition* 141 (2015): 127-37, at p. 130. Participants were professional philosophers and similarly educated non-philosophers and the effects persisted in the reflection condition across both groups.

[82] Anthony G. Greenwald, T. Andrew Poehlman, Eric Luis Uhlmann and Mahzarin R. Banaji, "Understanding and Using the Implicit Association Test: III. Meta-Analysis of Predictive Validity," *Journal of Personality and Social Psychology* 97 (2009): 17–41, at pp. 30-31

[83] Schroer and Schroer, "Two Potential Problems with Philosophical Intuitions."

relevant philosophical field] and, as a result, has unmuddled intuitions about those topics."[84] What, if anything, can be said in response to this objection? First, as Schroer and Schroer themselves acknowledge, the potential for misdirected judgements differs according to the conceptual complexity of the philosophical topic at hand. Thus the problem of misdirected judgements should be assessed on a case-by-case basis; it does not represent a general objection to employing public opinion. Moreover, there are a number of empirical strategies available to mitigate the problem of misdirected judgements: careful design of cases; inclusion of some explanatory material; and retrospective testing of participants, to identify and exclude those who have fallen prey to anticipated "muddles."[85] Of course, these methods do not fully guarantee against the problem of misdirected judgements. But to ask for such a guarantee is to demand too much, since our concern here is with the *relative* standing of the judgements of philosophers and non-philosophers. In this respect, Schroer and Schroer misdescribe the situation. They are right to claim that philosophers will typically have a better understanding of the relevant conceptual distinctions. However, they move too quickly from this premise to the conclusion that philosophers can thereby generate judgements fine-grained enough to target these distinctions. For example, take a political philosopher who makes a judgement about morally right action in a particular case. It is not obvious that she will be able to generate a separate and more specific judgement about the just action in this case – despite her understanding the general conceptual distinction between justice and moral rightness. Thus whilst the problem of misdirected judgements deserves serious attention, Schroer and Schroer do not demonstrate a clear asymmetry between the judgements of philosophers and non-philosophers in this regard.

Secondly, there is a potential problem of realism within popular attitudes. Being less familiar with the rules of the game when it comes to entertaining thought experiments, the public may be less willing to accept the parameters of the cases as they are given. In particular, they may be more likely

[84] Ibid, p. 1270.
[85] Ibid, pp. 1270-72.

to correct their responses for what they believe is reasonably likely to happen. For example, when considering how to deal with "push" versions of the trolley problem, they might (consciously or unconsciously) factor in the possibility that the heavy man will not, in fact, be enough to stop the trolley in its path.[86] I have anecdotal evidence of this kind of response to thought experiments on the part of non-philosophers, but as far as I am aware there are no systematic studies of its prevalence. Thus further research is needed. However, even if realism does turn out to be a problem, there are (as in the case of "muddled" judgements) potential means of correcting for it. For example, in their study of public responses to the trolley problem, Greene et al. direct respondents to "suspend disbelief" and exclude those who report being unable to do so. For the remaining respondents, they introduce a control for real world expectations.[87]

We can also make progress with the question of whose judgements are more "considered" by drawing on some emerging empirical research. Early work in experimental philosophy identified some troubling patterns in the responses of non-specialist audiences to philosophical thought experiments: judgements sometimes appear sensitive to the order in which the cases are presented or to the gender, cultural background or personality of the respondent.[88] Research has now moved on to compare patterns of judgement between philosophers and non-philosophers. For example, Schwitzgebel and Cushman report similar size case order effects on the judgements of professional philosophers and non-philosophers about a range of thought experiments in moral philosophy.[89] In a

[86] Joshua D. Greene, Fiery A. Cushman, Lisa E. Stewart, Kelly Lowenberg, Leigh E. Nystrom and Jonathan D. Cohen, "Pushing moral buttons: The interaction between personal force and intention in moral judgment," *Cognition* 111 (2009): 364-71, at p. 365.
[87] Ibid, pp. 366-7.
[88] On case order effects, see Stacy Swain, Joshua Alexander and Jonathan Weinberg, "The Instability of Philosophical Intuitions: Running Hot and Cold on Truetemp," *Philosophy and Phenomenological Research* 76 (2008): 138-55. On cultural variation, see Jonathan M. Weinberg, Shaun Nichols and Stephen Stich, "Normativity and Epistemic Intuitions," *Philosophical Topics* 29 (2001): 429-60. On personality effects, see Adam Feltz and Edward T. Cokely, "Do judgments about freedom and responsibility depend on who you are? Personality differences in intuitions about compatibilism and incompatibilism," *Consciousness and Cognition* 18 (2009): 342-50. These findings have been used by some to argue that cases should play a more limited role in philosophy. As noted in the introduction, it is not my intention to speak to this broader question, but rather to ask: assuming that case-based judgements do count, whose judgements should count?
[89] Eric Schwitzgebel and Fiery Cushman, "Expertise in Moral Reasoning? Order Effects on Moral Judgment in Professional Philosophers and Non-Philosophers," *Mind and Language* 27 (2012): 135-53. There is a question

follow-up study, they find that order and framing effects are not reduced among philosopher-participants reporting specialization in ethics and familiarity or expertise with the problems at hand.[90] Tobia et al. report that the responses of philosophers and non-philosophers do differ, but not in a way that gives greater credence to the philosophers' judgements. Specifically both groups show an actor-observer bias: their judgements vary according to whether they are described as the actor or the observer in a moral problem. However, the bias works in opposite directions for philosophers and non-philosophers.[91] Schulz and colleagues find that personality (specifically a facet of extraversion) exerts the same influence on judgements about free will and moral responsibility among those with different levels of expertise in this field.[92]

There is ongoing debate among experimental philosophers about the robustness and the proper interpretation of the initial findings of variability and instability in the responses of non-specialist audiences to philosophical cases.[93] Moreover, comparative work involving philosophers and non-philosophers is still in its early stages. However, the emerging results do not look promising for the claim that philosophers have superior expertise in responding to philosophical thought experiments. This is unsurprising in light of some key messages from the wider literature on expertise. Time spent practising an activity has been shown to be insufficient for the development of genuine expertise. Other factors are essential, in particular clear, timely and repeated feedback on task performance. As Weinberg et al. argue, such feedback is limited when it comes to the practice of forming

about how these findings relate to the problem of principle-bias discussed earlier: if philosophers are subject to such bias, presumably this ought to stabilize their intuitions? However, evidence of unstable judgements among philosophers does not demonstrate that principle-bias is not at work. At most it shows that, in some instances, such effects do not dominate other (also problematic) judgement-shaping factors.

[90] Schwitzgebel and Cushman, "Philosophers' biased judgments persist."

[91] Kevin Tobia, Wesley Buckwalter and Stephen Stich, "Moral intuitions: Are philosophers experts?" *Philosophical Psychology* 26 (2013): 629-38.

[92] Eric Schulz, Edward, T. Cokely and Adam Feltz, "Persistent bias in expert judgments about free will and moral responsibility: A test of the expertise defense," *Consciousness and Cognition* 20 (2011): 1722-31.

[93] Critiques of the experimentalists' methods and claims include: Simon Cullen, "Survey-Driven Romanticism," *Review of Philosophy and Psychology* 1 (2010): 275-96; Jennifer Nagel, "Intuitions and Experiments: A Defense of the Case Method in Epistemology," *Philosophy and Phenomenological Research* 85 (2012): 495-527; Robert L. Woolfolk, "Experimental Philosophy: A Methodological Critique," *Metaphilosophy* 44 (2013): 79-87.

judgements about philosophical cases.[94] In conclusion, the countervailing case for regarding philosophers' judgements as more considered is insufficient to outweigh the problem of principle-bias: the conclusion that the case-specific judgements of non-philosophers enjoy priority as an input into the process of RE stands.

6. Conclusion

Walter Lippman, in his famously critical discussion of the contribution of public opinion to politics, draws a distinction between insiders and outsiders, whose "relations to a problem are radically different. Only the insider can make decisions, not because he is inherently a better man but because he is so placed that he can understand and can act. The outsider is necessarily ignorant, usually irrelevant and often meddlesome, because he is trying to navigate the ship from dry land."[95] Lippman's contrast between insiders and outsiders is also helpful for thinking about the role of public opinion in moral and political philosophy. Specifically, I have argued that there is a form in which public opinion has value precisely *because* the public are outsiders to the terms of philosophical debate: public opinion better answers the demand of the method of RE for judgements about cases that are untainted by principled commitments. Thus, once we think carefully about what RE requires, there is a specific kind of ignorance that increases rather than decreases relevance. A case-based approach also serves to mitigate the methodological challenges that commonly confront efforts to integrate philosophical inquiry with the empirical investigation of popular attitudes.

This modestly public conception of RE was contrasted with some alternative approaches; in particular, the more comprehensively public method defended by Wolff and de-Shalit. Whilst they are right to advocate moving beyond an individual approach to RE, their approach, I argued, is overly ambitious about the work that public opinion can do within ethics and political philosophy. A wider

[94] Weinberg et al., "Are Philosophers Expert Intuiters?" pp. 340-42. See also Clarke, "Intuitions as Evidence," pp. 190-94.
[95] Walter Lippman, *The Phantom Public* (New Brunswick: Transaction Publishers, 2009), p. 140.

message of this paper is that it is a mistake to think about the task of integrating philosophy and public opinion as involving posing to the public all of the same questions that we ask ourselves as philosophers. Rather than treating the contributions of the philosopher and the public as symmetrical, we should draw on public opinion in ways that capture its distinctive value.[96]

Alice Baderin

Nuffield College, University of Oxford

alice.baderin@nuffield.ox.ac.uk

[96] An earlier version of this paper was presented at the Annual Conference of the Association for Legal and Social Philosophy, University of Leeds, 2014. I am grateful to the audience on that occasion, and to David Miller, Adam Swift and Jonathan Wolff, for their feedback. I thank Félix Krawatzek for help with German-language translation. The paper is based on work carried out during a DPhil funded by the Arts and Humanities Research Council. I gratefully acknowledge this support.