

ORIGINAL ARTICLE

Reasonable Disagreement and Metalinguistic Negotiation

Saranga Sudarshan 

Holsworthy, New South Wales, Australia

Correspondence

Saranga Sudarshan, 56 Lae Road,
Holsworthy, NSW 2173, Australia.
Email: nsaranga@gmail.com

Abstract

This paper defends a particular view of explaining reasonable disagreement: the Conceptual View. The Conceptual View is the idea that reasonable disagreements are caused by differences in the way reasonable people use concepts in a cognitive process to make moral and political judgements. But, that type of explanation is caught between either an explanatory weakness or an unparsimonious and potentially self-undermining theory of concepts. When faced with deep disagreements, theories on the Conceptual View either do not have the resources to explain them, or can only explain them by committing to a completely new theory of concepts where all moral and political concepts are a unique type of concept. This paper shows how the Conceptual View can avoid this dilemma by adopting what I dub the ‘Metalinguistic Strategy’ for explaining reasonable disagreement. The Metalinguistic Strategy uses recent innovations in the philosophy of language on metalinguistic negotiation to explain reasonable disagreements whether they be ordinary or deep disagreements as genuine disagreements whilst maintaining our ordinary ideas about concepts.

KEYWORDS

canonical dispute, metalinguistic negotiation, moral disagreement, political disagreement, reasonable disagreement

1 | INTRODUCTION

The question of what best explains reasonable disagreement is central to debates in political theory about the justification of political authority and coercive power. In those debates, political theorists advance theories about why political principles or rules ought to be obeyed and when it is permissible to coercively enforce them on the basis that their theories show how reasonable people who fundamentally disagree about moral and political matters can live in a shared social world.¹ A crucial part of arguing in this way involves theorists citing the best explanation of

¹See political liberals like Rawls (2005), Quong (2011), Nussbaum (2011), Hartley and Watson (2009, 2018), Larmore (1996, chap. 6; 1999), Gaus (2011, 2016) and Vallier (2019), and political realists like Williams (2005), Horton (2010), Rossi (2013), Hall (2015), Freyenhagen (2011) and Sleat (2013).

why reasonable people disagree about moral and political matters, and claiming that their theory shows how best to overcome those disagreements. As such, working out how political authority and the use of coercive power is justified hangs, at least in part, on what the best explanation of reasonable disagreement is.

But political theorists have no good options on this front. The contemporary debate about explaining reasonable disagreement is in a stalemate with explanations falling into three general views each with their own distinctive weaknesses. The simplest, and perhaps most popular, is the Imperfection View where what best explains reasonable disagreement is that people make a form of intellectual error in moral and political reasoning. On this view reasonable disagreement is caused either by some defect in reasoning,² or by people choosing the wrong type of reasoning for moral and political issues (e.g., Greene, 2013; Hare, 1981). But the Imperfection View faces two serious objections. Theories either rule out at least one aspect of what makes a disagreement reasonable, or simply fail to explain cases where reasonable people disagree without making any intellectual error.³ In response to these objections, theorists propose two alternatives.

On one alternative type of explanation, the Historical-Psychological View, what best explains reasonable disagreement is that people differ in their moral intuitions and the way their personal histories trigger these intuitions (e.g., Haidt, 2012; Turiel, 2002). On this view, the core idea is that the Imperfection View relies on a rationalist theory of moral psychology when a sentimentalist theory is better supported by empirical evidence. As such, the affective elements of cognition like moral intuitions are the foundation of moral and political judgement making and hence differences in those elements are what explain reasonable disagreement. But the problem for the Historical-Psychological View is that, as a number of theorists have pointed out, key aspects of the view (e.g., that non-affective judgement making is rare and that deliberative reasoning and intuitions are disparate cognitive acts)⁴ are not supported by the empirical evidence. Rather, the empirical evidence supports a more integrated model of moral and political judgement making where moral and political judgements are caused by a mix of emotional dispositions and deliberative reasoning.⁵

On another type of explanation, the Conceptual View, what best explains reasonable disagreement is that people possess and use their moral and political concepts in different ways when making moral and political judgements (e.g., Dworkin, 2011; Mason, 1993; McMahon, 2009). On this view, the core idea is that rather than adopting the pure rationalist moral psychology of the Imperfection View or the pure sentimentalist moral psychology of the Historical-Psychological View, we ought to adopt a more modest theory of moral psychology where both affective and non-affective cognition play a role in moral and political judgement making.⁶ On these theories, concepts whose content can be the result of both affective and non-affective cognition are the explanatorily significant elements. But the problem for the Conceptual View is that theorists are caught between a weakness in explanatory power and relying on a theory of concepts that is unparsimonious and self-undermining. They either cannot explain cases of deep disagreement as genuine disagreements, or, if they can explain them, require a belief in the existence of a novel and unique type of concept.

²This sort of explanation has an old pedigree in philosophy and is still widely popular. For instance, see Mason (1993, pp. 72–75) on how John Locke's theory of language could be used to formulate such an explanation. On more contemporary examples, see Enoch (2011, pp. 186–97, 207–14), Schafer-Landau (2003, chap. 9), Brink (1989, pp. 197–207), Wiggins (2006, pp. 366–67) and McDowell (1998, p. 162) for this sort of explanation. See Dworkin (2011, pp. 441–46) for an overview of how the latter two's views may be construed this way.

³See Mason (1993, p. 74) and Rawls (2005, p. 58) on the first sort of objection, and Kahane et al. (2012), Prinz (2016, pp. 57–60), Kahane (2016, pp. 291–92) on the second sort of objection.

⁴See Stanley et al. (2019) and Paxton et al. (2011, pp. 8–11) and Jacobson (2012, pp. 298–304) on the former, and Kennett and Gerrans (2016, pp. 76–82), Woodward (2016, p. 93), and Sauer (2012, pp. 266–70; 2015, pp. 161–63) on evidence against the latter.

⁵See Prinz (2016, pp. 67–68), Kennett and Fine (2009, pp. 88–91), Patterson et al. (2012), Young and Dungan (2012), and Suhler and Churchland (2011).

⁶See Nichols (2004) for such a theory and the empirical support for it.

Rather than trying to unravel and assess this dialectic as a whole, in this paper I will focus on a more narrow task. I will defend the viability of the Conceptual View and its middle way position between the Imperfection and Historical-Psychological Views. To salvage the Conceptual View, I argue that, by using recent innovations in the philosophy of language, it can explain cases of reasonable disagreement as genuine in one of two ways: as either canonical disputes or as metalinguistic negotiations. This avoids the charges of explanatory weakness and of an unparsimonious and self-undermining theory of concepts.

To that end, the paper proceeds as follows. In Section 2, I clarify the problem the Conceptual View faces and how it motivates the Metalinguistic Strategy. In Section 3, I lay out the core idea of the Metalinguistic Strategy, apply it to cases of reasonable disagreement that motivate it, and describe its advantages over its competitors. In Section 4, I consider and respond to two potential objections against the strategy.

2 | THE CONCEPTUAL VIEW AND DEEP DISAGREEMENT

Before going into the precise nature of the problem the Conceptual View faces, it is worth clarifying what reasonable disagreement is, and what it means to explain it. Reasonable disagreement is a state of affairs with four basic features that set it apart from other forms of disagreement. The first two of which mark it as an *intractable disagreement* about *moral and political issues*, and the last two mark it as *reasonable*.

The first feature is that the disagreement involves conflicting *moral or political* judgements. This means they involve conflicting judgements about the institutions and outcomes that justice, moral rightness or the good requires for a shared social world. By ‘institutions’ I mean the constitutions, ordinary statute laws and policies which Rawls (2005, pp. 11–12) referred to as the basic structure of a society. By ‘outcomes’ I mean the distribution of rights, opportunities and resources that those institutions produce. These institutions and outcomes are then evaluated according to people’s beliefs about justice, moral rightness or the good, which results in a judgement about the institutions and outcomes required. It is important to get clear about this because what I mean by ‘conflicting judgements about the institutions and outcomes that justice, moral rightness or the good requires’ is that people conflict in their evaluation of institutions and outcomes in virtue of their conflicting beliefs about what justice, moral rightness or the good requires.⁷ This means they make conflicting judgements in virtue of holding conflicting moral or political beliefs, and not in virtue of conflicting evaluations about what institutions and outcomes satisfy an agreed upon set of moral or political beliefs. The second feature is that the disagreement involves *intractably* conflicting judgements. This means that the disagreement persists despite the parties explaining their beliefs and their reasons for their judgement. As such it is not a disagreement that is a short-lived burst of conflict that occurs simply because people fail to spend enough time and effort trying to convince each other.

Those two features clarify reasonable disagreement as a species of moral and political disagreements in general. But moral and political disagreements in general are likely to occur for a variety of reasons that are uninteresting from the perspective of political theory. They could after all simply be caused by people being insincere, or from a lack of rationality. But what makes a disagreement distinctively *reasonable* is the sort of people they involve. This is what the third and fourth features describe.

The third feature of reasonable disagreement is that it involves people who both possess a minimal capacity for rationality. This means the disagreement does not occur because the people

⁷See Valentini (2013, pp. 183–87) for a helpful distinction between ‘thick disagreement about justice’ and ‘thin disagreement about justice’, which corresponds to the distinction I make here. Although Valentini ends up with a slightly different and more general formulation of reasonable disagreement, the core distinction matters both for her account and mine.

involved are unable to respond to moral reasons, make a coherent calculation of them or infer the reasons that are entailed by other ones. Rather, both parties to the disagreement have the capacity to make rational judgements. As such the disagreement cannot be resolved by one person deferring or being coerced by the other on the grounds they are unable to make rational judgements.⁸

Finally, the fourth feature is that the disagreement involves people who both have a minimal capacity for sincerely making judgements that others can agree to. This means the disagreement does not occur because one person is a contrarian who disagrees for the sake of disagreeing, or a fundamentalist who makes judgements that do not appropriately respond to the fact that they must try and live with others in a shared social world.⁹ Rather, both parties to the disagreement have the capacity to sincerely make judgements that their interlocutor could agree to. As such, the disagreement cannot be resolved by one person coercing another on the grounds the other is acting in bad faith or is amoral in some regard. Putting the four features together results in the following explanandum:

Reasonable Disagreement: A state of affairs of intractably conflicting moral or political judgements about a shared social world, made by at least two parties who both have a minimal capacity for rationality and a minimal capacity for sincerely making judgements that they think others can agree to.

This shows how reasonable disagreement differs in important ways from two sorts of disagreement discussed widely in philosophy. Firstly, reasonable disagreement is not faultless disagreement.¹⁰ Faultless disagreement is about a matter that is beyond rational conflict. Reasonable disagreement as described does not commit us to any such view about moral and political matters. Secondly, reasonable disagreement is not peer disagreement.¹¹ Peer disagreement occurs between parties that are actually epistemically equal in all regards. This would be an implausible picture of the sort of moral and political disagreements that matter for political theorists.¹² In sum, a reasonable disagreement is a disagreement between those of a certain moral and rational character and which is about moral and political matters that concern their shared social world.

At their core, theories that explain reasonable disagreement describe the facts that make the difference between a state of affairs of reasonable *disagreement* occurring and a state of affairs of reasonable *agreement* occurring. This description of difference makers explains how a case of reasonable disagreement is a genuine disagreement, and what causes the disagreement. As such, theories in the Conceptual View explain reasonable disagreement by citing facts about the role concepts play in people's judgement making. This involves proposing explanatory models that describe the cognitive process in which people use concepts to form conceptions and as a result make conflicting moral and political judgements. The idea being that concepts are the basic cognitive element that people use to form conceptions and make their moral and political judgements by weighing deliberative considerations associated with the concept. Conceptions are the general moral and political beliefs, typically principles that describe moral and political requirements, that people hold which they then use to make case-specific judgements about institutions and outcomes.¹³

⁸See, for similar construals of this aspect of reasonable disagreement, Rawls (2005, p. 52) on a 'capacity for a conception of the good', Gaus (2011, pp. 276–77) on a 'basic level of reasoning', and McMahon (2016, pp. 61–66) on 'reasonableness in the competence sense'.

⁹See Rawls (2005, p. 52), Gaus (2011, pp. 276, 288–92) and Carey (2018, pp. 51–59) for a similar minimalist sincerity condition as opposed to the far more demanding notion proposed by Quong (2011, pp. 265–73), and McMahon (2016, pp. 66–73) on 'reasonableness in the concession sense', which involves a similar idea of making judgements with a disposition 'to respond to perceived disparities of concession by making or seeking corrective concessions, provided that others are similarly disposed'.

¹⁰See Kölbel (2004) on the details of this type of disagreement.

¹¹See Kelly (2005) and Gutting (1982) on the details of peer disagreement and Peter (2013) and van Wietmarschen (2018) on its use in political liberalism.

¹²See Frances (2014, p. 166), King (2012), and Matheson (2014, pp. 320–28) on similar points about the irrelevance of peer disagreement.

¹³For this distinction between concepts and conceptions, see Rey (1983, 1985) for its early use in the philosophy of language and mind and Rawls (1999, pp. 5–6, 9; 2005, fn14) for its use in political philosophy.

One way in which theories on the Conceptual View cash out the role of concepts is to focus on how people use concepts to form conceptions. To get a sense of what that is like, a useful starting point is Rawls's (2005, pp. 54–58) explanation of reasonable disagreement through the 'burdens of judgement'.¹⁴ These burdens illustrate, according to Rawls, 'the many hazards involved in the correct (and conscientious) exercise of our powers of reason and judgement in the ordinary course of political life'. This explanation involves six sources of reasonable disagreement: (1) the complexity and conflicting nature of empirical evidence; (2) differences in the weight people afford deliberative considerations for making judgements; (3) differences in how we use concepts because of their inherent vagueness in certain hard cases; (4) differences in people's personal experiences affect how they weigh evidence and values; (5) differences in the kinds of deliberative considerations relevant to certain cases makes it difficult to weigh them against each other; and (6) social institutions are limited in the values they can embody which makes it difficult to weigh and prioritise some values over others.

Roughly put, the sources of disagreement all involve various ways that reasonable people weigh deliberative considerations (e.g., reasons, values, etc.) differently when making moral and political judgements. Of course, this general picture says little about how the sources interact or the precise causal process by which disagreement occurs. Given that, some philosophers have sought to develop it further.

For instance, Mason (1993, p. 58) argues that some moral and political concepts are *essentially contestable* such that they 'accredit a complex, valued achievement' and 'Different elements in this achievement may be weighted differently by different contestants'. According to Mason (1993, p. 59), this means that reasonable disagreement occurs because when people use their political concepts to form political beliefs, 'there will be a diversity of rational and non-rational causes of political belief formation'. The 'rational and non-rational' factors that affect how people form their political beliefs are, according to Mason (1993, pp. 15, 99–100), 'the reasons that people have for making the judgements' or in other words their deliberative considerations, and the 'psychological propensities or personal experiences' that affect how parties weigh those deliberative considerations.

Independently of Mason, McMahon (2009, pp. 3–4) has recently argued that what is relevant for understanding the cognitive process of how people use concepts to make moral and political judgements is how reasonable people learn to use moral and political terms. For McMahon (2009, pp. 55–58), learning how to use moral and political terms involves developing both an association with the features of a state of affairs that the moral and political terms apply to and the 'extrapolative dispositions' for using the term. The features of the state of affairs are a set of morally relevant deliberative considerations that a person learns to associate with a term's concept. The 'extrapolative dispositions' are dispositions individuals have for weighing their deliberative considerations and forming new beliefs or altering old ones when they encounter new states of affairs and have to use the term in question again. This will involve their extrapolative dispositions motivating them to extrapolate using a concept from what they have learned from their past personal experiences to the new case. For McMahon (2009, pp. 55–63, 78, 81), reasonable disagreement occurs because people will differ in the extrapolative dispositions they develop and the particular personal experiences they have where those dispositions are activated or developed further. This means they will weigh the deliberative considerations they use to form political beliefs slightly differently, and as a result form conflicting political beliefs.

Ultimately, Mason and McMahon's theories involve explanatory models that explain reasonable disagreement by describing how the burdens of judgement operate in a cognitive process where people use concepts to form conceptions and as a result make conflicting moral and political judgements. On these theories, reasonable disagreement occurs because people *use*

¹⁴See also Larmore (1996, chap. 7; 2008, chap. 8).

concepts differently to then form conflicting conceptions. To see how this works, consider the following case:

Abortion: Barry and Nora are discussing the laws concerning abortion that their society ought to enact. Barry, having grown up among poor single mothers, judges that their society ought to enact laws that permit abortion in the first trimester because he believes respecting women's bodily autonomy is what justice requires in this case. This is because respecting women's bodily autonomy outweighs the value of human life in a foetus. Nora, having grown up among poor women pressured to get abortions, judges that their society ought to outlaw abortion in the first trimester barring exceptional circumstances because she believes preserving the value of human life is what justice requires in this case. This is because the value of human life outweighs women's bodily autonomy.

This is a familiar sort of case that haunts much of our political life. Theories like Mason and McMahon's explain this case by citing the facts about the cognitive process involved in reasonable people making their political judgements. They cite a difference in Barry and Nora's psychological dispositions and personal experiences that lead them to weigh their deliberative considerations differently and form conflicting conceptions of justice. Mason would cite differences in Barry and Nora's 'psychological propensities and personal experience' with respect to abortion and pregnancies, whilst McMahon would cite their 'extrapolative dispositions' and their learning experiences for the term 'justice', 'autonomy', and perhaps 'murder'. In both models, the point is that Barry and Nora's different personal experiences cause them to be disposed to weigh the value of life and bodily autonomy differently, and, as a result, form conflicting conceptions of justice. This causes them to make conflicting judgements about the institutions and outcomes related to abortion that justice requires.

But, theories like Mason and McMahon's face a problem when presented with cases of reasonable disagreement that are also deep disagreements. Deep disagreements are cases where people exchange mutually intelligible reasons and make persistently conflicting judgements as a proxy for a systematic conflict of 'worldviews', 'perspectives' or 'fundamental commitments'.¹⁵ For an example of a reasonable disagreement that is a deep disagreement consider the following:

Nationalisation: Bryan and Elizabeth are discussing the economic structure their society ought to have. Elizabeth judges that their society ought to nationalise, at the very least, some key industries because a society's productive capacity being for the mutual benefit of all is what justice requires. Bryan judges that their society ought not to nationalise any industries because protecting people's natural moral right to their body and private property is what justice requires. This is because he believes natural moral rights are what matter, benefiting everyone with society's productive capacity is irrelevant and has nothing to do with what justice requires. On the contrary, Elizabeth believes materially benefiting people is what matters, natural rights to private property are irrelevant and have nothing to do with what justice requires.

This is a case of a reasonable disagreement that is a deep disagreement because Bryan and Elizabeth's disagreement about the economic structure justice requires is a proxy for a deeper systematic conflict in their fundamental views about how one ought to think, deliberate and form beliefs about what justice requires. Unlike Barry and Nora in *Abortion*, they do not

¹⁵See Ranalli (2021, pp. 984–986), Hazlett (2014, pp. 12–13), Kappel (2021, p. 1039), and Fogelin (1985) for an overview of deep disagreement in epistemology, and Gaus (145–46; 2011, pp. xv–xvii, 2, 38–42; 2016, 2017, 2018, p. 647), Knoll (2018, pp. 24, 39–43) and Muldoon (2016, pp. 1–4) in political philosophy.

disagree simply in virtue of the content of their beliefs about what justice requires. Rather, what Bryan and Elizabeth's disagreement about the just economic structure hinges on is their conflicting views about the considerations that ought to be used to make judgements about what justice requires. This rings true of many contemporary political disagreements. For instance, consider disagreements over whether the United Kingdom ought to leave or remain in the European Union, whether transgender people ought to have access to single-sex spaces, or whether healthcare ought to be distributed by private healthcare insurance markets. All these disagreements seem to hinge not simply on the beliefs or judgements reasonable people hold about particular issues, but on their entire view of how reasonable people ought to go about deliberating when making moral and political judgements about those issues.

Cases like *Nationalisation* pose a problem for Mason and McMahon's theories. They are cases of disagreement which Mason and McMahon's theories cannot explain. This is because their theories do not have the resources to explain the conflict in views *about* the deliberative considerations that ought to be used to make moral and political judgements. Their entire focus, in extending Rawls's burdens of judgement, was to describe how people *use* concepts differently to form conflicting moral and political beliefs by describing how people weighed the deliberative considerations differently. But cases of deep disagreement outstrip the resources of that explanation. Citing facts about how reasonable people weigh their deliberative considerations cannot explain conflicts about which considerations ought to be weighed.¹⁶

One way the Conceptual View tries to avoid the problem of deep disagreement is by proposing another way to cash out the role of concepts in people's moral and political making, namely through how people possess concepts as well as merely how they use them. These theories involve explanatory models that cite how the possession of a concept in addition to its use affects how people form conceptions and as a result make moral and political judgements. The core idea is that facts about how people possess concepts will be facts about the deliberative considerations a concept provides when forming conceptions. Given that deep disagreements involve a conflict over what the relevant deliberative considerations are when making moral and political judgements about some particular issue, facts about how people possess concepts would explain this conflict.

Ronald Dworkin's theory of interpretive concepts offers an explanation of this sort, the core idea being that moral and political concepts are interpretive concepts and therefore possessed and used in such a way that they lead to reasonable disagreements rather than agreements. As such, Dworkin (2011, pp. 159–60) makes a distinction between 'criterial' and 'natural-kind' concepts on the one hand and 'interpretive' on the other. Criterial and natural-kind concepts are characterised by the fact that people 'do not share a concept of either kind unless they would accept a decisive test', as such 'Genuine disagreement about application is ruled out once all pertinent facts are agreed upon'. For Dworkin, explaining reasonable disagreement and vindicating the intuition that cases like *Nationalisation* are genuine requires recognising that moral and political concepts are not like this. They are interpretive.

Interpretive concepts have three distinctive features.¹⁷ The first is that as moral and political concepts, interpretive concepts describe values (Dworkin, 2011, p. 6). These values involve an interpretation of what is of value and ought to be furthered in a particular social practice of using that concept. To that end, the content of interpretive concepts involves values that are the result of an interpretation of the concept itself. The second feature is that to use an interpretive concept is to engage in conceptual interpretation. It is to identify what is of value or disvalue in a given social practice, by interpreting the purpose of identifying what is of value or disvalue in

¹⁶One thought might be that these theories could simply shift down a level and describe how people use concepts differently in their 'disagreement about how to use a concept'. The problem here is that concepts are a basic and fundamental part of cognition. There are no further cognitive elements that underlie it that can be cited as being used or weighed differently. Thank you to an anonymous reviewer for pointing out this potential response.

¹⁷See also Plunkett and Sundell (2013b, pp. 251–52) for this broad understanding of Dworkin's theory.

the social practice that we are engaged in. Given this, Dworkin (2011, pp. 162–63) argues that conceptual interpretation involves the possession of other interpretive concepts, namely those that describe the array of values that one has to weigh up when interpreting the initial concept. This is what Dworkin (2011, p. 154) means when he says that conceptual interpretation is ‘pervasively holistic’. The use of interpretive concepts relies on the interpretation of a network of concepts we possess.¹⁸ This means that what is identified as valuable or not valuable in a shared social practice will rely on people using other interpretive concepts, and so any single act of interpretation is guided by how it fits with the entire web of interpretations. The third feature is that reasonable people can be said to possess and use the same concept despite differing in their interpretation of what they are doing with the concept if they all agree on certain paradigm uses of the concept. For Dworkin (2011, pp. 160–61), sufficient agreement about paradigm uses of a concept (i.e., particular judgements) will ensure that everyone is using the same interpretive concept. This will be the case even if people interpret the content of the concept so differently they hold vastly different beliefs about what the appropriate values are when engaging in the shared social practice of using that concept.

With interpretive concepts so understood, Dworkin (2011, pp. 150–51, 162) argues that the facts that make the difference between reasonable disagreement and reasonable agreement are an individual’s experiences and interpretive dispositions that affect their conceptual interpretations across their network of interpretive concepts. This is because the concepts that feature in reasonable disagreements are moral and political concepts, and those concepts are interpretive concepts. This means that ordinary cases of reasonable disagreement like *Abortion* are explained by differences in reasonable people’s interpretive dispositions and personal experiences causing them to interpret a concept like *JUSTICE* differently such that they weigh a shared set of values in different ways. In these cases, reasonable people share the interpretation of their network of interpretive concepts to a sufficient degree that they identify the same values as the best interpretation of the aim and purpose of the social practice of using *JUSTICE* as a concept. But, their psychological dispositions and personal experiences cause them to weigh these values differently such that they form conflicting conceptions of justice and make conflicting political judgements according to them. Cases of deep disagreement like *Nationalisation* are explained by differences in people’s interpretation of their network of interpretive concepts such that they identify different values as the best interpretation of the aim and purpose of the social practice of using *JUSTICE* as a concept to form conceptions of justice. This is what explains how a reasonable disagreement can turn on a conflict about what the appropriate inputs to deliberation are for forming conceptions and judgements about moral and political issues.

It seems, then, that the Conceptual View has a way to escape the charge of explanatory weakness. Dworkin’s theory of interpretive concepts can explain both ordinary cases of reasonable disagreement and those that are deep disagreements. But this improvement comes at a heavy theoretical cost which opens up two new problems. The first problem is that it is unparsimonious.¹⁹ Salvaging the Conceptual View through Dworkin’s theory requires a commitment to the existence of an entirely new type of concept, interpretive concepts, and a non-standard view of non-interpretive concepts. Dworkin’s theory requires the Conceptual View to break with the best contemporary accounts of how people think and talk and enlarge our theoretical resources purely for the purpose of explaining reasonable disagreement.

The second problem is that the wide scope of its claims about concepts is self-undermining. Dworkin’s theory claims *all* moral and political concepts are interpretive concepts and not

¹⁸Dworkin (2011, pp. 162–63) does warn that there is a non-foundationalist circularity to his theory. But he argues that the circularity is justified by his stronger thesis about the independence and holistic unity of true interpretations of all our values. See Winter (2016) and Knight (2006) for discussion of this view. I leave this question about the holism of value aside as it is irrelevant to how Dworkin’s theory explains reasonable disagreement.

¹⁹See Plunkett and Sundell (2013b, pp. 253–55) on this charge in context of Dworkin’s theory and the more narrow issue of the concept of legal disagreement.

merely those involved in a particular reasonable disagreement. This potentially undermines the explanation it can provide of reasonable disagreement in at least two ways. It is self-undermining in one way because reasonableness itself will end up being an interpretive matter. Therefore, whether a given case of reasonable disagreement is explained can itself be the subject of reasonable disagreement. Someone can always interpret ‘reasonable’ differently. This disagreement would then require some new explanation. It is self-undermining in another way because when reasonable people in deep disagreements try to communicate why they make their judgements, Dworkin’s theory entails that they will descend into a regression of interpretation. After all, they cannot communicate why they interpret the content of a certain moral or political concept differently without interpreting all the other moral and political concepts that it is connected to, which in turn requires more interpretation.

In sum, when faced with reasonable disagreements that are deep disagreements, the Conceptual View faces a dilemma. Theories like Mason and McMahon’s cannot explain deep disagreements, whilst theories like Dworkin’s are committed to a theory of concepts that is unparsimonious and potentially self-undermining.

Salvaging the Conceptual View requires, then, a way for it to explain deep disagreements without introducing new types of concepts. This is what I take up in the rest of this paper. In the next section I propose that, by making use of recent innovations in the philosophy of language on normative and evaluative disagreement, specifically on metalinguistic negotiations, we can explain reasonable disagreements that are deep disagreements without requiring Dworkin’s theory of concepts.

3 | REASONABLE DISAGREEMENTS AND THE METALINGUISTIC STRATEGY

3.1 | Canonical disputes and metalinguistic negotiations

The core idea of what I propose is that we can explain why reasonable disagreements occur by recognising that reasonable disagreements are not all genuine in the same way. That is to say, by making use of Plunkett and Sundell’s innovative analyses of normative and evaluative disagreements, reasonable disagreements are sometimes ‘canonical disputes’ and sometimes they are a particular form of non-canonical dispute, a ‘metalinguistic negotiation’.²⁰ Before seeing how such an explanation works, it is worth getting clear on some of the moving parts in this account and exploring precisely what canonical disputes and metalinguistic negotiations are.

The first moving part is that on Plunkett and Sundell’s account, meaning (understood as some linguistically encoded information for a word that allows us to use the word to refer to things) is taken to be closely connected to concepts (understood as the basic building blocks of thought which are the mental equivalent of words). They are taken to be connected such that the meaning of a word is determined by the concept it is used to express and a single word can express different concepts just as it can mean different things depending on the context.

The second moving part is that a genuine disagreement must involve some conflict in mental content.²¹ Roughly put, this means that for two people to disagree their view of the truth or correctness of some belief, judgement, plan, intention or some mental content is rationally incompatible. In short, there must be some mental content that one person takes as true or correct and the other takes to be false or incorrect.

²⁰See also Plunkett and Sundell (2013b, pp. 256–73).

²¹See Frances (2014, chap. 1) and Gibbard (2003, chap. 4) for this basic way of thinking about genuine disagreement.

The upshot of these two moving parts is that they sit well with the account of reasonable disagreement given so far. We have seen how the Conceptual View also takes the same sort of close connection between word meaning and concepts as Plunkett and Sundell. In addition, we have seen how reasonable disagreement involves conflicts over moral and political judgements. This gives us a good indication that Plunkett and Sundell's account can be applied to reasonable disagreements.

With the moving parts of Plunkett and Sundell's account and its appropriateness for reasonable disagreements clarified, we are now in a position to deal with the particulars of their view. According to Plunkett and Sundell, when disagreements are canonical disputes, they are genuine in virtue of a conflict in what speakers literally express. This explanation involves two claims. The first claim is that the disagreement hinges on a conflict in some mental content (e.g., beliefs, judgements, plans, intentions) that speakers *literally* express, or would literally express given the opportunity. The second claim is that the speakers likely mean the same thing by their words because otherwise they would not conflict in what they literally express. They would be expressing mental contents with different truth-conditions or conditions of correctness such that the mental contents would not be inconsistent.

To get a sense of how such an explanation works consider the following sort of everyday disagreement:

Salt: Alice and Beth are cooking soup and are in a debate about whether there is any salt in the soup. Alice judges the soup is unsalted as she has been watching the whole time and no one put any salt in it. But Beth judges there is salt in it because she saw someone put it in.

Although *Salt* is a fairly simple case, the best explanation using Plunkett and Sundell's account is that it is a canonical dispute. We have good evidence, given what they say, that it hinges on some content that they literally express or would express to each other. Alice and Beth would, given what they claim to have seen, literally express incompatible judgements about the salt in the soup. We also have good evidence they likely mean the same thing by their words, because given, what they are doing and what they claim to have seen, they are likely to use the word 'salt' in a systematically similar way across various contexts. There is no indication that either are likely to make drastic shifts in the way they use the word 'salt'.

But Plunkett and Sundell (2013a, pp. 11–13; 2013b, pp. 247–48) argue that when it comes to normative and evaluative disagreements, not all cases are like *Salt*. Rather, some disagreements are best explained as 'non-canonical disputes', the idea being that some cases of disagreement are genuine not in virtue of a conflict in the mental content that speakers *literally* express. Rather, they are genuine in virtue of a conflict in the mental content that speakers *pragmatically* express. Of these non-canonical disputes, the type that I believe is ultimately helpful for explaining reasonable disagreements are 'metalinguistic negotiations'. A metalinguistic negotiation is, as Plunkett and Sundell (2015, pp. 837–51; 2013a, pp. 13–18; 2013b, pp. 256–66) have argued, a disagreement that is *non-canonical* in virtue of consisting of a conflict about what a word ought to mean. This means the content that is pragmatically expressed is a 'metalinguistic belief' about what the meaning of a word ought to be.²² This also means that disagreeing parties mean different things by, at least, one of their words. This is because it is the use of words with different meanings that pragmatically expresses the belief about what the word ought to mean.

To get a better sense of how all this works consider the following case:

Spicy Soup: Oscar and Callie are cooking soup for a party and are in a heated debate about its spiciness. They both taste the soup and Oscar, with the taste palates of the party guests

²²See also Chalmers (2011, pp. 522–23) on these implicit metalinguistic beliefs.

in mind, judges that the soup is spicy. But Callie, with her long experience tasting chillies in mind, disagrees and judges the soup not spicy at all.²³

The point of *Spicy Soup* is that we have a strong intuition that it is a genuine disagreement. But if it is read as a canonical dispute it would not be. This is because it does not seem to hinge on what Oscar and Callie believe about the soup, but rather what they think the threshold for spiciness should be. To solve this, we can read it as a metalinguistic negotiation. This explanation involves two claims. The first claim is that Oscar and Callie's disagreement involves a conflict in the mental content that speakers *pragmatically* express, namely their metalinguistic beliefs about what the meaning of 'spicy' ought to be. The second claim is that they must mean different things by 'spicy'. It is, after all, the use of 'spicy' with different meanings that pragmatically, rather than literally, expresses their metalinguistic belief about what it ought to mean.

But what evidence is there for these claims? One piece of evidence for the first claim, again following Plunkett and Sundell (2015, p. 847; 2013a, pp. 19–20), is that disagreeing parties are likely to carry on disagreeing even when they both agree about what their terms currently mean in each other's expressions, in the community at large or according to some authoritative third party. This clearly seems to be the case in *Spicy Soup*. Oscar and Callie seem disposed to carry on their disagreement even if they could agree on what 'spicy' means to other third parties or how it is defined in a recipe book. They are each committed to their meaning because many normative consequences could flow from whether it is resolved on Oscar or Callie's behalf. Amongst other things, it could affect whether the soup ought to be served, whether it ought to be praised more or less by the party guests, or even whether Oscar and Callie ought to cook in future.

With the second claim, again following Plunkett and Sundell (2013a, p. 15), one piece of evidence for it is that the disagreeing parties are disposed to use their terms in systematically different ways across varying contexts. Or as Plunkett (2015, p. 847) says whether 'speakers are disposed to systematically use a term in divergent ways in the same (non-defective) conditions'. This matters because the meaning of a word depends, at least in part, on our patterns of using that word. To that extent Oscar and Callie do seem disposed to systematically use 'spicy' in divergent ways in the process of cooking the soup. By this I mean they are not disposed to suddenly use the term in the same way to refer to the same things as 'spicy'. Therefore, if word use is a guide to its meaning, it seems plausible to conclude that they mean different things by 'spicy'.

3.2 | Explaining reasonable disagreements

Now, with all that in mind, I propose extending the strategy of explaining a disagreement as either a canonical dispute or a metalinguistic negotiation to reasonable disagreements, specifically, by explaining ordinary reasonable disagreements as canonical disputes and reasonable disagreements that are deep disagreements as metalinguistic negotiations.²⁴

To understand how we can explain cases of reasonable disagreement as canonical disputes consider again the case of *Abortion*:

²³This is a slightly modified case that Plunkett and Sundell (2013a, pp. 14–15) use to illustrate the power of reading disagreements as metalinguistic negotiation.

²⁴Note, Plunkett and Sundell believe a metalinguistic analysis generally (not merely metalinguistic negotiation) can be plausibly extended to all kinds of normative and evaluative disagreements (2013a, pp. 7, 18–25; 2013b, pp. 265–66). But they do not touch upon how it can be used in the context of reasonable disagreements and in particular deep disagreements. One reason to think it can is the way the distinction between 'content' and 'character' allows metalinguistic negotiations to involve both gradable terms like 'spicy' and moral terms which might have more fixed meanings (Plunkett, 2015, pp. 837–38; Plunkett and Sundell, 2013a, pp. 15–16).

Abortion: Barry and Nora are discussing the laws concerning abortion that their society ought to enact. Barry, having grown up among poor single mothers, judges that their society ought to enact laws that permit abortion in the first trimester because he believes respecting women's bodily autonomy is what justice requires in this case. This is because respecting women's bodily autonomy outweighs the value of human life in a foetus. Nora, having grown up among poor women pressured to get abortions, judges that their society ought to outlaw abortion in the first trimester barring exceptional circumstances because she believes preserving the value of human life is what justice requires in this case. This is because the value of human life outweighs women's bodily autonomy.

Given Plunkett and Sundell's account from Section 3.1, I propose the best way to understand how a disagreement like *Abortion* is a genuine disagreement is to read it as a canonical dispute. This would, as we saw in the previous section, involve claiming that Barry and Nora conflict over what they literally express, and that they likely mean the same things by their words. In the case of *Abortion*, there is clear evidence for making both claims. The main piece of evidence for the first claim is that what Barry and Nora literally express, or more accurately what they *would* literally express, are each of their beliefs about the abortion laws justice requires. These beliefs, Barry's that abortion should be permitted because respecting women's bodily autonomy is what justice requires and Nora's that abortion should not be permitted because preserving the value of human life is what justice requires, involve conflicting mental content. They cannot both be true.

As for the claim that Barry and Nora likely mean the same things by their words, the main piece of evidence for this is that, the reason Barry and Nora believe what they do is that they both think that the value of bodily autonomy and human life are relevant for determining what justice requires and as such for using the term 'justice' correctly. They both think justice is a matter of how bodily autonomy and human life are weighed up. This is what allows them to express logically incompatible mental content. If, for instance, Nora believed that abortion should not be permitted because the only thing relevant for justice is to respect the souls of other creatures, and that a foetus has a soul, this would not be logically inconsistent with Barry's belief. The truth-conditions for Barry and Nora's beliefs about justice would be entirely different.²⁵

They are also disposed to use the word 'justice' in a systematically similar way across various contexts. It is only their experiences growing up that push them to weigh the considerations relevant for justice differently and reach conflicting beliefs about the justice of abortion. However, in other contexts, like compulsory vaccines or patient consent laws, they would use the term 'justice' the same way.

This provides the Conceptual View with the adequate resources to then explain why a reasonable disagreement like *Abortion* occurs. It can easily cite the differences in Barry and Nora's personal experiences or psychological disposition that make them use their concept of justice in slightly different ways to form conflicting conceptions of justice, these conceptions being their beliefs about 'the right distribution of rights, opportunities and resources amongst people, institutions and social systems.'²⁶ When Barry and Nora use their conflicting beliefs about justice to make case-specific judgements, they end up making intractably conflicting political judgements which comprise *Abortion*.

²⁵Of course, in such a case, the disagreement would be a good candidate for a metalinguistic negotiation. But as *Abortion* is presented here, it is not.

²⁶This explication of a *conception* of justice summarises what I take Rawls (1999, pp. 5–6, 9, 54) to be talking about in various places as the 'proper distribution of the benefits and burdens of social cooperation', 'the appropriate distributive shares' and the point of his own conception of justice being to describe the right distribution of the chief primary goods—'rights, liberties, and opportunities, and income and wealth'—that the basic structure of society—the political constitution and the principal economic and social arrangements—can dispense.

So far so good. But what about reasonable disagreements that are deep disagreements? As I have said, I propose we should explain them by reading them as metalinguistic negotiations. To understand better what that means consider the type of case that motivated it, namely a reasonable disagreement that is a deep disagreement:

Nationalisation: Bryan and Elizabeth are discussing the economic structure their society ought to have. Elizabeth judges that their society ought to nationalise, at the very least, some key industries because a society's productive capacity being for the mutual benefit of all is what justice requires. Bryan judges that their society ought not to nationalise any industries because protecting people's natural moral right to their body and private property is what justice requires. This is because he believes natural moral rights are what matter, benefiting everyone with society's productive capacity is irrelevant and has nothing to do with what justice requires. On the contrary, Elizabeth believes materially benefiting people is what matters, natural rights to private property are irrelevant and have nothing to do with what justice requires.

Given Plunkett and Sundell's account, I believe the best way to explain a case like *Nationalisation* as a genuine disagreement is to read it as a metalinguistic negotiation. In a case of deep disagreement like *Nationalisation*, this involves the following two claims. The first claim is that Bryan and Elizabeth's disagreement is genuine in virtue of involving a conflict in their metalinguistic beliefs about what the meaning of 'justice' ought to be. This is the belief they *pragmatically* express (or would express if given the opportunity), by *literally* expressing their judgements about what justice requires. The second claim is that Bryan and Elizabeth mean different things by 'justice'. Making those two claims allows us, despite such a case not turning on what reasonable people literally express about what justice requires, to understand how it is still in fact a genuine case of disagreement.²⁷

This strategy seems to hold because the reasons for making those two claims in a case of deep disagreement like *Nationalisation* are the same as they were for *Spicy Soup*. The evidence for the first claim is that it seems Bryan and Elizabeth are disposed to carry on disagreeing even when they both agree on what 'justice' currently means in the community at large or according to some third party. After all, they are not pointing to dictionaries or articles in philosophy journals to defend their moral and political judgements. At least one of them thinks those things are irrelevant. As such, the beliefs they literally express are not inconsistent as the truth-conditions for those beliefs as determined by what they take to be relevant for determining the use of 'justice' as different.²⁸ But despite this, they are likely to persist in their disagreement because they recognise that using 'justice' with the meaning that their interlocutor uses has significant normative consequences for them. It matters what states of affairs are described as 'just' because it will affect what normative demands are placed on people and what social arrangements will be enforced with the use of coercive political power.

For the second claim, it seems that Bryan and Elizabeth are disposed to use 'justice' in systematically divergent ways whenever they need to make judgements pertaining to matters of justice. This is because, as we have already established, they disagree about what the

²⁷See Ball (2020) for someone who is sceptical of metalinguistic negotiation as a way to explain deep disagreements. This is because understanding the participants as advancing a view about what the meaning of a word ought to be, either (1) entails that people are being unreasonable in trying to advance the first-order issue (e.g., the economic institutions and outcomes justice requires) by asserting a view about what the meaning of a word ought to be, or (2) fails to make sense of the reaction that a reasonable person has to have to contest the first-order claim rather than contest the view about what the meaning of a word ought to be. I am not sceptical in this way. I think Ball's analysis goes wrong because he overlooks the way pragmatic expression is key to metalinguistic negotiation, and that he places an epistemic standard of reasonableness on the participants that is either irrelevant to political theory or a standard political theorist need not accept.

²⁸Notice this is a crucial difference to *Abortion* where the literal expressions of the interlocutors do conflict because each agrees on the truth-conditions of their beliefs.

deliberative considerations ought to be when making moral and political judgements, and what deliberative considerations are used will determine, in some sense, how people use their moral terms across various contexts.²⁹ The idea being that when people do not even agree about what deliberative considerations to use, they are disposed, across various contexts, to form conflicting beliefs about what their moral terms apply to. If they then make their judgements accordingly, they will *ipso facto* use their terms in systematically divergent ways. To that end, if patterns of word use are a guide to a word's meaning, then Bryan and Elizabeth mean different things by 'justice'. As such, we have evidence that *Nationalisation* hinges on a pragmatically expressed conflict about what 'justice' ought to mean and therefore is a metalinguistic negotiation. All this shows how reasonable disagreements that are ordinary disagreements and those that are deep disagreements can be cases of genuine disagreement. The former are canonical disputes and the latter are metalinguistic negotiations.

The strategy of explaining a case like *Nationalisation* as a metalinguistic negotiation provides the Conceptual View with the resources to then explain why it occurs. It can cite, as it did with the case of *Abortion*, the differences in a speaker's personal experiences and psychological dispositions. But this time it cites those differences as causing Bryan and Elizabeth to make conflicting *metalinguistic* beliefs about 'justice' which are beliefs about what the word 'justice' ought to mean. Since the meaning of a word has a close connection to the concept because the meaning of a word is determined by the concept that a person possesses, conflicting metalinguistic beliefs in effect mean that Bryan and Elizabeth possess different concepts. This will mean that when they use those different concepts to deliberate and form conceptions of justice, they will weigh up completely different considerations. For Bryan this will be something like 'natural moral rights to their body and private property are the morally relevant considerations for justice'. For Elizabeth it will be something like 'productive capacity being for the mutual benefit of all are the morally relevant considerations for justice'. In real-world cases, what these conceptual contents actually are will be more complex and nuanced and subject to actual empirical investigation. But for now, it suffices to show how a conflict over their metalinguistic beliefs leads them to disagree.

3.3 | Comparative advantages and the vindication of the Conceptual View

The last two subsections have explored Plunkett and Sundell's account of normative and evaluative disagreements, and then applied this account to reasonable disagreements. This showed how we can vindicate the Conceptual View as an explanation of reasonable disagreements. The strategy of explaining reasonable disagreement as either canonical disputes or metalinguistic negotiations has two clear comparative advantages over extant theories in the Conceptual View. Recall, the problem with the extant theories is that they faced a dilemma. On one horn, theories like Mason and McMahon's are explanatorily weak because they do not have the resources to explain deep disagreements. The theories focus on how people *use* concepts differently to make conflicting political judgements. As such they have nothing to say about what causes deep disagreements where the substantive normative conflict is about something other than the political judgements people make.

But using Plunkett and Sundell's account of disagreements has proven more explanatorily powerful than theories like Mason and McMahon's by actually explaining deep disagreements. It does this because it allows for the possibility that reasonable disagreements may be canonical disputes or metalinguistic negotiations. This means they can sometimes hinge on pragmatically

²⁹Notice again this is a crucial difference to *Abortion* where the interlocutors agree on the deliberative considerations that determine the correct usage of 'justice'.

expressed conflicting beliefs about what a moral term ought to mean and sometimes hinge on literally expressed conflicting moral and political beliefs.

On the other horn of the dilemma, Dworkin's theory of interpretive concepts is the only extant theory that can explain deep disagreements, but is unparsimonious and potentially self-undermining. It is unparsimonious because it requires breaking with the best contemporary accounts of how people think and talk to commit oneself to the existence of an entirely new type of concept. It is potentially self-undermining in at least two ways because it claims that *all* moral and political concepts are interpretive concepts and not merely those involved in a particular reasonable disagreement. This undermines the explanation it can provide in one way because reasonableness itself will end up being an interpretive matter and therefore whether Dworkin's theory explains reasonable disagreement will itself be the subject of reasonable disagreement. This matters because reasonableness consists of a set of normative features, namely the capacity for a minimal capacity for rationality and a minimal capacity to make sincere judgements that others can agree to. These involve interlocutors who believe they should respond sincerely to the reasons at hand and make judgements that they think others could agree to. If reasonableness itself was interpretive, we would not be able to settle on a case of reasonable disagreement to explain in the first place.

In another way, when reasonable people in deep disagreements try to communicate why they make their judgements, Dworkin's theory entails that they will descend into a regression of interpretation. This is because any act of communication would require interpreting more concepts at every turn.

Making use of Plunkett and Sundell's account avoids all these issues because it relies on a far more parsimonious view of concepts and how they work. Rather than posit an entirely new type of concept, it merely posits a new type of genuine disagreement. This leaves concepts as they are and allows for a theory to make use of the best empirical data on how they work.

4 | OBJECTIONS

Taking stock for a moment, so far I have argued for the Metalinguistic Strategy as a way to salvage the Conceptual View of explaining reasonable disagreement on the basis of two lines of argument. First, that it has more explanatory power than certain extant theories because it can explain cases of reasonable disagreement that are both ordinary disagreements and deep disagreements. Second, that it is better than other extant theories because it does not rely on any unparsimonious commitments or potentially self-undermining view about concepts. I want to now consider two objections that might be raised against the first line argument, namely that employing the idea of metalinguistic negotiations has inadvertent consequences for reasonable disagreement. The first concerns the pointlessness of verbal disputes. The second concerns topic discontinuity.

4.1 | Pointless verbal disputes

I have argued that one advantage of applying Plunkett and Sundell's account to reasonable disagreements is that it allows for the explanation of reasonable disagreements that are deep disagreements as genuine disagreements. This is because it allows for the reading of such disagreements as metalinguistic negotiations that hinge on a pragmatically expressed conflict about what the meaning of a moral term ought to be. An objection one might have to this, following Chalmers (2011, pp. 522–25), is that it makes deep disagreements into pointless verbal disputes, the core idea being that, even if there is a sense in which deep disagreements, when read as metalinguistic negotiations, are genuine in virtue of hinging on a conflict about what concept ought to be used, they are pointless to the first-order practical matter that is literally expressed by reasonable people's political judgements. And, as Chalmers (2011, p. 525) says,

this sort of pointlessness is a heuristic guide to the presence of verbal disputes. As such, using metalinguistic negotiation to vindicate deep disagreements as genuine disagreements was in vain. They are pointless verbal disputes anyway and pointless disputes are no help for political theorists who hope to show reasonable disagreements were worth explaining.

To see how this objection works, consider *Nationalisation* again. In that case, the thought would be that, when it is read as a metalinguistic negotiation, it turns out to hinge not on the first-order disagreement we thought it did. It does not hinge on the economic system Bryan and Elizabeth believe justice requires. Rather, it hinges on a conflict about what the concept of JUSTICE ought to be. But insofar as this generates the first-order disagreement, it is pointless to it because all it involves is a conflict about what considerations are relevant for deliberating about justice. And, resolving that dispute will resolve the first-order disagreement. But this resolution would not come about because Bryan and Elizabeth settled something about justice. Rather, it would come about because they settle on which considerations they ought to be weighing up to form beliefs about what justice requires.

But, I submit, this objection goes wrong in at least one important way. It is plainly wrong to conclude that metalinguistic negotiations are pointless with respect to the first-order moral and political issues that reasonable people are disagreeing about. The resolution of a metalinguistic negotiation about what concept ought to be used has significant normative consequences when they involve the kinds of concepts that usually feature in reasonable disagreements. Reasonable disagreements, as should be plain from examples like *Abortion* and *Nationalisation* can potentially involve concepts like JUSTICE, MORAL GOODNESS, MORALLY RIGHT, MORAL PERSON and many other moral concepts. These are by their very nature normative concepts and therefore if people diverge over them and decide to resolve the divergence, they will have serious and pervasive normative consequences. For instance, that some states of affairs are picked out as required by JUSTICE is typically taken to justify realising it with the use of coercive political power. As such, people will be forced to comply with laws that enforce a state of affairs that counts as just according to their interlocutor's concept. One can imagine how this will play out in a similar way for a whole host of reasonable disagreements that feature moral and political concepts. To that end, deep disagreements read as metalinguistic negotiations are not pointless, and as such not necessarily or heuristically verbal disputes.

For a non-political example, consider a reasonable disagreement that is centred on a topic in interpersonal morality, 'the good life'. Let us suppose that it is a metalinguistic negotiation and as Chalmers supposes parties resolve to use the term 'good life' to mean 'lives spent satisfying selfish desires' rather than to use 'good life' to mean 'lives spent helping others'. Clearly this will have profound normative differences for these individuals. It will affect what they see as demanded of them by morality and therefore how they should lead their lives.

4.2 | Topic discontinuity

Another objection against the Metalinguistic Strategy might be that, even if we accept that understanding deep disagreements as metalinguistic negotiations it does not entail they are pointless verbal disputes. Rather it might entail they involve a form of topic discontinuity. That is to say that the model shows that deep disagreements are cases where reasonable people are trying to change the topic of their discussion, rather than engaging in a sustained disagreement about the same topic.

The objection begins with the thought that concepts play a central role in picking out the 'topic' or general subject matter of thought.³⁰ For instance, when someone is deciding to buy a

³⁰See Sawyer (2018a, 2018b, 2020), Schroeter and Schroeter (2014) and Cappelen (2018, chap. 9) for an overview of this type of view of the relation between concepts and topics.

house and reflecting about what they ought to do, a number of thoughts will typically run through their head. They will think about various things relating to their finances, various things relating to the condition of the house, perhaps what others think about the house or their finances as well. But all through this, their thoughts centre on a particular topic picked out by a concept, or more likely a cluster of concepts, relating to 'purchasing a house'. This will be the same for people engaged in a discussion or disagreement. A sameness in concepts will guarantee a sameness in topic.

But, so the objection goes, if the strategy I have argued for is right, that reasonable disagreements like *Nationalisation* are metalinguistic negotiations, then it shows that the people in those disagreements are not actually talking about the same topic. This is because as the Metalinguistic Strategy explains those cases, the disagreeing parties are each making a judgement according to their divergent concepts. This is not a pointless verbal dispute, but it does indicate that the disagreeing parties are trying to have a discussion centred on different topics, namely topics individuated by the concepts they are each individually using. This would be to say that in *Nationalisation*, Bryan is trying to discuss what justice requires as it pertains to his concept of JUSTICE, but Elizabeth is trying to discuss what justice requires as it pertains to her concept. But, neither of them realise that they are each trying to change the topic and so there is no sustained disagreement on a single topic of discussion.

But this objection is unwarranted because the Metalinguistic Strategy is fundamentally neutral with respect to a wide variety of views on topic continuity. For instance, it is compatible with the sort of view argued for by Schroeter and Schroeter (2014, pp. 12–16), and Sawyer (2018a, pp. 10–15; 2018b, pp. 13–21; 2020, pp. 385–90) where concept identity guarantees topic continuity. They argue topic sameness depends on people using the same concepts, which is in turn understood with people sharing a 'tradition' for fixing the content of that particular concept. This tradition is supposed to include all the various representational and non-representational ways that people's thoughts are related to the world. The Metalinguistic Strategy is compatible with such a view because it recognises that people must at least share some core conceptual content pertaining to the role of the concept in thought, and because it is neutral on precisely which facts fix that core content. As such, the theory would say that if topic continuity depends on concept identity, we should merely look at the core content of a concept, and when people overlap on that it indicates they 'diverge' in their concepts in the sense of putting forth candidates for the same concept. Since the Metalinguistic Strategy is not committed to any particular story about how or which facts fix this content, it is entirely compatible with a shared tradition of content fixing as Schroeter and Schroeter, and Sawyer propose.

Beyond views that cash out topic sameness through concept identity, the Metalinguistic Strategy is also compatible with a range of other views as well. For instance, it is entirely compatible with a view like Cappelen's (2018, pp. 107–8) that addresses topic continuity with the idea that it is a pre-theoretic notion. This is because it does not use topic continuity as a theoretical notion or commit to any broader metasemantic theory about entities other than concepts, meaning, extension and intension. As such, it is neutral on what pre-theoretic notions can be posited to make sense of topic continuity. This means it is compatible with the solution offered by Cappelen, which makes use of the notion of 'samesaying' as a pre-theoretic coarse-grained notion independent of a word's meaning or concept.

Finally, the Metalinguistic Strategy is also compatible with a purely pragmatic view of topic continuity of the sort argued for by Roberts (2012). For instance, the topic for a reasonable disagreement could be supplied by 'questions under discussion'. These are questions that set the function of the discourse between the parties. The initial question is then something like 'What institutions and outcomes does justice require?'. This would then elicit conflicting answers from reasonable people. This in turn gives rise to further questions under discussion that are ever more general to finally answer the initial question. On this view what seems like discontinuity of topic is really the work of pragmatic mechanisms in the disagreement to ultimately answer

the initial question. The use of Plunkett and Sundell's account is entirely compatible with a view like this because it has taken no position on the pragmatic mechanisms involved in a disagreement beyond those involved in metalinguistic negotiations. To that end, it is entirely compatible with additional pragmatic mechanisms if need be.

5 | CONCLUSION

In this paper I have defended the viability of the Conceptual View by proposing that Plunkett and Sundell's account of normative and evaluative disagreements be applied to explaining reasonable disagreements. I did this first by exploring how this account, which I dub the Metalinguistic Strategy, can explain disagreements as either canonical disputes or as metalinguistic negotiations. I argued this strategy allows us to read cases of reasonable disagreement that are both ordinary and deep disagreements as genuine. I then applied this distinction to the cases of reasonable disagreement that motivated the strategy to begin with. I then argued that this strategy has two clear comparative advantages that vindicate the Conceptual View. Firstly, that by being able to explain deep disagreements, it is more explanatorily powerful than extant theories that claim to explain reasonable disagreement. Secondly, that it maintains our ordinary view of concepts unlike Dworkin's theory. I then considered and responded to two objections that might be put against the strategy.

ACKNOWLEDGEMENTS

Thank you to Derek Ball, Benjamin Sachs, Gerald Gaus, Kevin Scharp, Clotilde Torregrossa, Quentin Pharr, Lisa Bastian, Ethan Landes, Mirela Fus and Ravi Thakral for specific comments on parts of this paper, and to the entire Arche Conceptual Engineering Seminar Group between 2016 and 2020 at the University of St Andrews, and the audience at the Arche Reading Party 2018, for helpful comments, discussion and encouragement along the way.

ORCID

Saranga Sudarshan  <https://orcid.org/0000-0002-0750-0177>

REFERENCES

- Ball, D. (2020) Revisionary analysis without meaning change (or, could women be analytically oppressed?). In: Burgess, A., Cappelen, H. & Plunkett, D. (Eds.) *Conceptual engineering and conceptual ethics*. Oxford: Oxford University Press.
- Brink, D. (1989) *Moral realism and the foundations of ethics*. Cambridge: Cambridge University Press.
- Cappelen, H. (2018) *Fixing language: an essay on conceptual engineering*. Oxford: Oxford University Press.
- Carey, B. (2018) Public reason – honesty, not sincerity. *The Journal of Political Philosophy*, 26(1), 47–64.
- Chalmers, D. (2011) Verbal disputes. *Philosophical Review*, 11(4), 515–566.
- Dworkin, R. (2011) *Justice for hedgehogs*. Cambridge, MA: Harvard University Press.
- Enoch, D. (2011) *Taking morality seriously: a defense of robust realism*. Oxford: Oxford University Press.
- Fogelin, R. (1985) The logic of deep disagreements. *Informal Logic*, 7(1), 1–8.
- Frances, B. (2014) *Disagreement*. Cambridge: Polity Press.
- Freyenhagen, F. (2011) Taking reasonable pluralism seriously: an internal critique of political liberalism. *Politics, Philosophy & Economics*, 10(3), 323–342.
- Gaus, G. (2011) *The order of public reason: a theory of freedom and morality in a diverse and bounded world*. Cambridge: Cambridge University Press.
- Gaus, G. (2016) *The tyranny of the ideal: justice in a diverse society*. Princeton, NJ: Princeton University Press.
- Gaus, G. (2017) Is public reason a normalization project? Deep diversity and the open society. *Social Philosophy Today*, 33(1), 27–52.
- Gaus, G. (2018) The complexity of a diverse moral order. *The Georgetown Journal of Law & Public Policy*, 16(1), 645–680.
- Gibbard, A. (2003) *Thinking how to live*. Cambridge, MA: Harvard University Press.
- Greene, J. (2013) *Moral tribes: emotion, reason, and the gap between us and them*. Harmondsworth: The Penguin Press.

- Gutting, G. (1982) *Religious belief and religious skepticism*. Notre Dame, IN: University of Notre Dame Press.
- Haidt, J. (2012) *The righteous mind: why good people are divided by politics and religion*. Harmondsworth: The Penguin Press.
- Hall, E. (2015) Bernard Williams and the basic legitimation demand: a defence. *Political Studies*, 63(2), 466–480.
- Hare, R.M. (1981) *Moral thinking: its levels, method, and point*. Oxford: Clarendon Press.
- Hartley, C. & Watson, L. (2009) Feminism, religion, and shared reasons: a defense of exclusive public reason. *Law and Philosophy*, 28(5), 493–536.
- Hartley, C. & Watson, L. (2018) *Equal citizenship and public reason: a feminist political liberalism*. Oxford: Oxford University Press.
- Hazlett, A. (2014) Entitlement and mutually recognized reasonable disagreement. *Episteme*, 11(1), 1–25.
- Horton, J. (2010) Realism, liberal moralism and a political theory of *modus vivendi*. *European Journal of Political Theory*, 9(4), 431–448.
- Jacobson, D. (2012) Moral dumbfounding and moral stupefaction. In: Timmons, M. (Ed.) *Oxford studies in normative ethics: volume 2*. Oxford: Oxford University Press.
- Kahane, G. (2016) Is, ought, and the brain. In: Liao, M. (Ed.) *Moral brains: the neuroscience of morality*. Oxford: Oxford University Press.
- Kahane, G., Wiech, K., Shackel, N., Farias, M., Savulescu, J. & Tracey, I. (2012) The neural basis of intuitive and counterintuitive moral judgment. *Social Cognitive and Affective Neuroscience*, 7(4), 393–402.
- Kappel, K. (2021) Higher order evidence and deep disagreement. *Topoi*, 40, 1039–1050.
- Kelly, T. (2005) The epistemic significance of disagreement. In: Gendler, T.S. & Hawthorne, J. (Eds.) *Oxford studies in epistemology: volume 1*. Oxford: Oxford University Press.
- Kennett, J. & Fine, C. (2009) 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(1), 77–96.
- Kennett, J. & Gerrans, P. (2016) The rationalist delusion?: A post hoc investigation. In: Liao, M. (Ed.) *Moral brains: the neuroscience of morality*. Oxford: Oxford University Press.
- King, N. (2012) Disagreement: What's the problem? Or a good peer is hard to find. *Philosophy and Phenomenological Research*, 85(2), 249–272.
- Knight, C. (2006) Justice for foxes. *Law and Philosophy*, 34(6), 633–659.
- Knoll, M. (2018) Deep disagreements on social and political justice: their meta-ethical relevance and the need for a new research perspective. In: Knoll, M., Snyder, S. & Şimsek, N. (Eds.) *New perspectives on distributive justice*. Berlin: De Gruyter.
- Kölbel, M. (2004) III—faultless disagreement. *Proceedings of the Aristotelian Society*, 104(1), 53–73. Available from: <https://doi.org/10.1111/j.0066-7373.2004.00081.x>
- Larmore, C. (1996) *The morals of modernity*. Cambridge: Cambridge University Press.
- Larmore, C. (1999) The moral basis of political liberalism. *The Journal of Philosophy*, 96(12), 599–625.
- Larmore, C. (2008) *The autonomy of morality*. Cambridge: Cambridge University Press.
- Mason, A. (1993) *Explaining political disagreement*. Cambridge: Cambridge University Press.
- Matheson, J. (2014) Disagreement: idealized and everyday. In: Matheson, J. & Vitz, R. (Eds.) *The ethics of belief: individual and social*. Oxford: Oxford University Press.
- McDowell, J. (1998) *Mind, value, and reality*. Cambridge, MA: Harvard University Press.
- McMahon, C. (2009) *Reasonable disagreement: a theory of political morality*. Cambridge: Cambridge University Press.
- McMahon, C. (2016) *Reasonableness and fairness: a historical theory*. Cambridge: Cambridge University Press.
- Muldoon, R. (2016) *Social contract theory for a diverse world: beyond tolerance*. London: Routledge.
- Nichols, S. (2004) *Sentimental rules: on the natural foundations of moral judgment*. Oxford: Oxford University Press.
- Nussbaum, M.C. (2011) Perfectionist liberalism and political liberalism. *Philosophy & Public Affairs*, 39(1), 3–45.
- Patterson, R., Rothstein, J. & Barbey, A.K. (2012) Reasoning, cognitive control, and moral intuition. *Frontiers in Integrative Neuroscience*, 6(114), 1–8.
- Paxton, J.M., Ungar, L. & Greene, J. (2011) Reflection and reasoning in moral judgment. *Cognitive Science*, 36(1), 1–15.
- Peter, F. (2013) Epistemic foundations of political liberalism. *Journal of Moral Philosophy*, 10(5), 598–620.
- Plunkett, D. (2015) Which concepts should we use?: Metalinguistic negotiations and the methodology of philosophy. *Inquiry*, 58(7–8), 828–874.
- Plunkett, D. & Sundell, T. (2013a) Disagreement and the semantics of normative and evaluative terms. *Philosophers' Imprint*, 13(23), 1–37.
- Plunkett, D. & Sundell, T. (2013b) Dworkin's interpretivism and the pragmatics of legal disputes. *Legal Theory*, 19(3), 242–281.
- Prinz, J. (2016) Sentimentalism and the moral brain. In: Liao, M. (Ed.) *Moral brains: the neuroscience of morality*. Oxford: Oxford University Press.
- Quong, J. (2011) *Liberalism without perfection*. Oxford: Oxford University Press.
- Ranalli, C. (2021) What is deep disagreement? *Topoi*, 40, 983–998.
- Rawls, J. (1999) *A theory of justice*, revised edition. Cambridge, MA: Belknap Press of Harvard University Press.

- Rawls, J. (2005) *Political liberalism*, expanded edition. New York: Columbia University Press.
- Rey, G. (1983) Concepts and stereotypes. *Cognition*, 15(1–3), 237–262.
- Rey, G. (1985) Concepts and conceptions: a reply to Smith, Medin and Rips. *Cognition*, 19(3), 297–303.
- Roberts, C. (2012) Information structure in discourse: towards an integrated formal theory of pragmatics. *Semantics and Pragmatics*, 5(1), 1–69.
- Rossi, E. (2013) Consensus, compromise, justice and legitimacy. *Critical Review of International Social and Political Philosophy*, 16(4), 557–572.
- Sauer, H. (2012) Educated intuitions. Automaticity and rationality in moral judgement. *Philosophical Explorations*, 15(3), 255–275.
- Sauer, H. (2015) Can't we all disagree more constructively? Moral foundations, moral reasoning, and political disagreement. *Neuroethics*, 8(2), 153–169.
- Sawyer, S. (2018a) Subjective externalism. *Theoria*, 84(1), 4–22.
- Sawyer, S. (2018b) The importance of concepts. *Proceedings of the Aristotelian Society*, 118(2), 1–21.
- Sawyer, S. (2020) Talk and thought. In: Burgess, A., Cappelen, H. & Plunkett, D. (Eds.) *Conceptual engineering and conceptual ethics*. Oxford: Oxford University Press.
- Schafer-Landau, R. (2003) *Moral realism: a defence*. Oxford: Oxford University Press.
- Schroeter, L. & Schroeter, F. (2014) Normative concepts: a connectedness model. *Philosophers' Imprint*, 14(25), 1–26.
- Sleat, M. (2013) *Liberal realism: a realist theory of liberal politics*. Manchester: Manchester University Press.
- Stanley, M., Yin, S. & Sinnott-Armstrong, W. (2019) A reason-based explanation for moral dumbfounding. *Judgement and Decision Making*, 14(2), 120–129.
- Suhler, C.L. & Churchland, P. (2011) Can innate, modular “foundations” explain morality? Challenges for Haidt's moral foundations theory. *Journal of Cognitive Neuroscience*, 23(9), 2103–2116.
- Turiel, E. (2002) *The culture of morality: social development, context, and conflict*. Cambridge: Cambridge University Press.
- Valentini, L. (2013) Justice, disagreement and democracy. *British Journal of Political Science*, 43(1), 177–199.
- Vallier, K. (2019) *Must politics be war?: Restoring our trust in the open society*. Oxford: Oxford University Press.
- van Wietmarschen, H. (2018) Reasonable citizens and epistemic peers: a skeptical problem for political liberalism. *The Journal of Political Philosophy*, 26(4), 486–507.
- Wiggins, D. (2006) *Ethics: twelve lectures on the philosophy of morality*. Cambridge, MA: Harvard University Press.
- Williams, B. (2005) *In the beginning was the deed: realism and moralism in political argument*. Princeton, NJ: Princeton University Press.
- Winter, J. (2016) Justice for hedgehogs, conceptual authenticity for foxes: Ronald Dworkin on value conflicts. *Res Publica*, 22(4), 463–479.
- Woodward, J. (2016) Emotion versus cognition in moral decision-making: a dubious dichotomy. In: Liao, M. (Ed.) *Moral brains: the neuroscience of morality*. Oxford: Oxford University Press.
- Young, L. & Dungan, J. (2012) Where in the brain is morality? Everywhere and maybe nowhere. *Social Neuroscience*, 7(1), 1–10.

How to cite this article: Sudarshan, S. (2023) Reasonable Disagreement and Metalinguistic Negotiation. *Theoria*, 1–20. Available from: <https://doi.org/10.1111/theo.12454>