It's common sense – you don't need to believe to disagree!

Miklós Kürthy, Graham Bex-Priestley & Yonatan Shemmer

Published online: 13 Apr 2023.

Submit your article to this journal ⍐

Article views: 212

View related articles ⍐

View Crossmark data ⍐

Routledge
Taylor & Francis Group

It's common sense – you don't need to believe to disagree!

Miklós Kürthy [a], Graham Bex-Priestley [b] and Yonatan Shemmer [a]

aDepartment of Philosophy, University of Sheffield, Sheffield, UK; bSchool of Philosophy, Religion and History of Science, University of Leeds, Leeds, UK

ABSTRACT

It is often assumed that disagreement only occurs when there is a clash (e.g., inconsistency) between beliefs. In the philosophical literature, this "narrow" view has sometimes been considered the obvious, intuitively correct view. In this paper, we argue that it should not be. We have conducted two preregistered studies gauging English speakers' intuitions about whether there is disagreement in a case where the parties have non-clashing beliefs and clashing intentions. Our results suggest that common intuitions tell against the default view. Ordinary speakers describe clashes of intentions as disagreements, suggesting that the ordinary concept of disagreement is "wide" in that it extends beyond beliefs.

1. Introduction

Can two people disagree without having conflicting beliefs? The philosophical literature is divided on this question. The traditional view is that disagreement is always grounded in a clash between beliefs (henceforth: narrow view), while more recently, a number of researchers have defended the view that disagreement is possible in virtue of a clash between other attitudes such as desires, intentions, emotions or plans (henceforth: wide view). It is evident that there are apparent cases of non-belief disagreement. However, the narrow view is that such apparent disagreement is merely derivative; that it is always an expression of disagreement in the underlying beliefs. There are very few explicit statements of the narrow view,[1] let alone arguments in its support. We suspect this is not because its proponents are unconvinced of their own position, but rather because the truth of that position seems so obvious to them that they find no need to argue for it. Time and again articles and books about disagreement set out the debates

---

CONTACT Miklós Kürthy m.kurthy@sheffield.ac.uk Department of Philosophy, University of Sheffield, 45 Victoria Street, Sheffield S3 7QB, UK

All three authors contributed equally to this project.

they engage in while completely ignoring the possibility of non-belief disagreement.[2]

In other papers (Bex-Priestley & Shemmer, 2017; Shemmer & Bex-Priestley, 2021) we have considered at length the philosophical arguments in favor of the wide view. But what are the arguments for the narrow view? As we pointed out above, its proponents do not put forth any arguments, perhaps because the view appears to fit well with our intuitions (as suggested by MacFarlane, 2007, p. 22). But is the narrow view in fact in line with common intuition? Our paper is dedicated to finding this out. As we shall see, the empirical studies discussed below (section 4) indicate that common intuition favors the wide view. Consequently, we believe that the burden of proof lies with advocates of the narrow view to provide arguments in its favor. In this essay we present the studies and discuss their implications.

2. Some examples

Claudius believed that the sun revolved around the earth. Nic thought that this wasn't so; the earth, he believed, revolved around the sun. Theirs was a straightforward disagreement in beliefs.

Cato the Elder was in favor of an immediate and unwavering campaign against Carthage. Many senators disagreed; they were for peace and quiet. If Cato and the senators indeed disagreed, their disagreement was, at least on the face of it, a disagreement in non-belief attitudes.

Other non-belief attitudes are mentioned as candidates for disagreement. People sometimes speak of disagreement in taste, in desire, in credence, in intention, and in emotion. For many particular instances of a clash between non-belief attitudes one could plausibly deny that they are cases of disagreement. If Ben likes Vanilla Ice-Cream and Jerry likes Cookies and Cream, it is far from clear that they disagree.

Nevertheless, the rejection of particular examples as cases of disagreement leaves open the question of whether there are any instances of genuine non-belief disagreement. As we will see in the next section, a theoretically attractive strategy is to understand every apparent case of non-belief disagreement as a derivative case of belief disagreement and therefore as a case in which the disagreement is not fundamentally a disagreement between the non-belief attitudes.

3. A short history of disagreement

The survey article about disagreement found in the Stanford Encyclopedia of Philosophy describes disagreement in the following way: "two individuals disagree about a proposition when they adopt different doxastic attitudes toward the same proposition" (Frances & Matheson, 2019, §1). The article

distinguishes between belief-disagreement and action-disagreement, but contends that: "Despite this distinction, we can achieve some simplicity and uniformity by construing disagreements over what to do as disagreements over what to believe. We do it this way: if we disagree over whether to do action X, we are disagreeing over the truth of the claim 'We should do X'" (Frances & Matheson, 2019, §2).[3]

The approach of the SEP article is an instance of what we have called above the narrow view. According to this view, any apparent disagreement over actions, intentions, emotions, desires, plans, etc. is, or is the outward expression of, disagreement about whether or not it is true that one should adopt the attitude in question.[4] In other words, it is, or it is the expression of, a disagreement in beliefs. The narrow view has its advantages. It is simple, it provides a unified account of disagreement, and it may fit naturally with those cases of disagreement that most readily come to mind when we think about the notion.

Nonetheless, mainly in recent years, some philosophers have attempted to take seriously the idea that non-belief disagreement is not only possible but is in fact as fundamental as belief disagreement. This new approach has its origins in the need of non-cognitivists in metaethics to analyze disagreement in moral judgment in terms of conative attitudes but has now spread well beyond the realm of non-cognitivism. Notable examples are: Bex-Priestley and Shemmer (2017); Björnsson et al. (2010); Blackburn (1998); Dreier (1999); Egan (2014); Finlay (2014); Gibbard (2003); Huvenes (2012, 2014); MacFarlane (2014); Marques (2015); Marques and García-Carpintero (2014); McKenna (2014); Priest (2006); Richard (2015); Ridge (2013); Stevenson (1944); Sundell (2010); Worsnip (2019).[5] The range of analyses of the notion of disagreement in the "wide" camp is quite broad, but the camp is unified in thinking that there is nothing pre-theoretically objectionable about the idea of fundamental non-belief disagreement.

The question is therefore: what is our pre-theoretical understanding of the idea of disagreement? Both the proponents of the narrow view and the proponents of the wide view are likely to claim that they have the right grasp of our pre-theoretical notion. But such claims, given the theoretical commitments of their would-be utterers, must be regarded with suspicion. For this reason, an empirical study into the common pre-philosophical understanding of the concept of disagreement is of value.

Before we present our own research, a few words about the existing empirical literature. Several studies, mainly in the last fifteen years, have explored attitudes toward people in disagreement. Notable examples are Goodwin and Darley (2008, 2010, 2012); Nichols (2004); Sarkissian et al. (2011); Theriault et al. (2017); Wright et al. (2013); Wright et al. (2014).[6] Broadly speaking this body of work aimed to assess whether the folk have implicit objectivist metaethical views.

One of the recurring research methods used in these studies is what Beebe calls "the disagreement paradigm".[7] Respondents are presented with situations of moral disagreement and are then asked to make metaethical judgments about the disagreeing parties. For example, they are asked to decide whether one of those who are party to the disagreement is mistaken, or whether there is a fact of the matter regarding whether one of the judgments is correct. In the background of that work lies an understanding of disagreement as an activity. On that understanding, by definition, two people disagree if they reject each other's claim by saying "no". Thus, if I say, "Banks are corrupt" and you say "No, banks prevent the river from overflowing", then on this understanding we disagree, whether or not the content of our statements, our beliefs, or any other set of attitudes of ours, are in conflict.

While these studies rely on judgments about situations of disagreement and are therefore in the broad ballpark of our own research, there is no direct connection between this body of work and our project. This is because we, unlike the studies mentioned above, do not stipulate a notion of disagreement. Rather, our research is part of a large theoretical literature[8] interested in how best to characterize the concept of disagreement. And while there is no consensus in that literature about what this characterization should be, most parties to the debate reject the notion of disagreement as an activity.[9]

An exception to the disconnect between the existing empirical research and the question we are interested in here, is a paper by Khoo and Knobe titled "Moral Disagreement and Moral Semantics" (Khoo & Knobe, 2018). Khoo and Knobe are interested in the possibility of disagreement between utterers of conflicting moral statements that is not grounded in the conflict between the semantic content of these statements, or as they put it "disagreement without exclusionary content". Their vignettes present members of distant cultures who make conflicting moral statements; for example, statements about the actions of Dylan, a random stabber. The one says "Dylan didn't do anything morally wrong" and the other says "No, Dylan did do something morally wrong". Khoo and Knobe provide convincing empirical evidence that participants think that the utterers disagree even when they do not think that the utterers' statements have exclusionary content. They proceed to offer a possible explanation. According to their explanatory hypothesis participants judge disagreement without exclusionary content when they see utterers as making conflicting proposals for updating the accepted norms of the conversational context (more on that below).

Let us briefly consider the similarities and differences between their approach and ours.

(1) While Khoo and Knobe share with other empirical researchers the understanding of "disagreement" as an activity,[10] their research partly resembles our own both in aim and methodology. Both their studies and ours wish to investigate which situations are seen by participants as situations of disagreement. In particular, in one of the studies, they carry out this investigation independently of their earlier stipulative usage of the term.[11]

(2) Our targets are related but different. Khoo and Knobe are interested in disagreement that is not grounded in the semantic content of the statements of those who disagree. We are interested in disagreement that is not grounded in the content of any of the beliefs of those who disagree, neither the beliefs expressed by their statements nor any other beliefs whose presence can explain the disagreement in question.

(3) Their research is narrower in scope. They only establish the possibility of *moral* disagreement without exclusionary content, and possibly without the *moral* beliefs expressed by that content. Our aim is to establish the possibility of *generic (moral or non-moral)* disagreement without belief.

(4) Khoo and Knobe's findings are impressive given their own philosophical target. However, they have limited relevance to the possibility of disagreement without belief, since an objector to disagreement without belief might claim that while the contents of the statements expressing the moral conflict are not exclusionary, the utterers of these statements must have (other) implicit beliefs whose content is exclusionary in order for them to disagree. In particular, an objector to the possibility of disagreement without belief might appeal to Khoo and Knobe's positive account to buttress her objection. According to Khoo and Knobe, the best way to account for disagreement without exclusionary content is to adopt a certain form of contextualism, according to which the truth of moral claims depends on the standards/norms accepted in the context of utterance, together with the view that in certain situations, e.g., when the utterers belong to distant cultures, these standards might be indeterminate. As a result, it is possible for people from very different cultures to make moral claims that are the negation of each other without either of the claims being false. They then locate the disagreement in the incompatibility between the speech acts of the utterers. The statements made by the parties to the disagreement, say Khoo and Knobe, function as proposals to update the conversational context so as to change the common ground between speakers with regards to the accepted moral standards. The proposals to update the conversational context are

incompatible when it is impossible for there to be contexts where both proposals are accepted. However, the objector to disagreement without belief is unlikely to be satisfied with this account.[12] Incompatible proposals are not necessarily cases of disagreement. Consider Serena and Roger who converse during their tennis match. Roger: "I propose you give up and let me win the game"; Serena: "No, I propose *you* give up and let me win". In most instances the proposals of Serena and Roger do not constitute disagreement, but rather mere expressions of diverging desires. They would constitute disagreement, according to the objector, only if there were underlying beliefs, implicit or explicit, with exclusionary content; e.g., and quite implausibly in this case, if each one of the players believed that it is the right thing for the other player to let them win the game. Thus, while the content of the utterances studied by Khoo and Knobe might not be exclusionary, the conclusion that disagreement exists in situations of moral conflict that are the subject of their studies must depend, thinks the objector, on the attribution of conflicting implicit beliefs to the utterers. For example, one utterer might believe that the norms that ought to be accepted as common ground in the conversational context are ones which support random stabbing whereas the other utterer might believe that the norms that ought to be accepted as common ground in the conversational context are ones which reject random stabbing. Without (implicit) beliefs with such exclusionary content, the objector will claim, there is no reason to see the different proposals as cases of disagreement; rather, we should see them in the same way we see the proposals of the tennis players above.

(5) Since our own experiments aim to establish the possibility of disagreement without belief, they are designed to block this very objection.

(6) Khoo and Knobe's positive theory of disagreement without exclusionary content is disunified in an important respect: it gives a different account of the nature of disagreement in cases where it is grounded in exclusionary content and in cases where it is grounded in conflicting speech acts. Elsewhere (Bex-Priestley & Shemmer, 2017, p. 190) we argued that a unified account of disagreement must be favored over a non-unified account.

In what comes below, we report the results of two studies that were designed to test whether the application of the term "disagreement" to clashes of non-belief attitudes is indeed unnatural (section 4), and we

discuss the philosophical implications of the combined empirical findings (section 5).

4. Empirical studies

4.1. General overview

The aim of our study was to establish whether competent users of English consider it natural to apply the term "disagreement" to certain clashes in non-belief attitudes that could not be understood (and are not understood by those speakers) as expressions of disagreement in belief.[13] Our assumption is that a natural inclination to apply the term "disagreement" to certain circumstances is a good guide to the user's implicit understanding of the concept signified by this term. We have presented subjects with a vignette describing two individuals with differing intentions under circumstances where it is hard to ascribe to them differing beliefs about what these intentions should be. We have then endeavored to establish whether subjects would readily apply the term "disagreement" to the relation between those individuals.

Our working assumption was that if people think of disagreement as a relation which must always express a clash between beliefs, then the overwhelming majority of participants would not describe the individuals depicted in the vignette as disagreeing with one another. If, on the other hand, disagreement is not understood as being exclusively such a relation, but instead might also be a clash of other propositional attitudes such as intentions, then participants would be ready to describe the situation as one involving disagreement.

4.2. Study 1

4.2.1. Method

We originally recruited 559 native speakers of English via Amazon Mechanical Turk, of whom 202 were eliminated due to failed control questions (see below). Our final sample thus consisted of 357 participants (203 male, 153 female, 1 other, M_{age} = 37.23, SD = 10.92). The study was pre-registered (see osf.io/2wyd6). We presented participants with the following vignette:

> Buba and Kiki are friends. They live in a remote community in which being indecisive is seen as a terrible vice. Members of this community have developed strategies in order not to be indecisive. Whenever they face a choice between options that seem equally good to them, they pick one of the options arbitrarily and then they form an intention to act on that option—without changing their view that the options are equally good.

> Buba and Kiki have decided to spend the evening together. The options available to them are to go to a movie or to go to a restaurant. Since, to both of them, these options seem like an equally good way to spend the evening together (they both think that there isn't a best way to spend the evening together), each one of them resorts to the strategy for avoiding indecisiveness. Buba arbitrarily forms the intention to go to a movie and Kiki arbitrarily forms the intention to go to a restaurant.
>
> "I am only willing to go to a movie", says Buba. "I am only willing to go to a restaurant", says Kiki.

Having read the story, participants were asked three main questions about (a) whether Buba and Kiki disagree (*disagreement probe*), (b) whether Buba and Kiki have changed their minds after implementing their community's "strategy" (*view-change probe*), and (c) whether Buba and Kiki end up having different intentions at the end of the story (*intention probe*).[14] The order of these questions varied as a function of condition (see the next section).

Finally, we included three control questions inquiring into some details of the story (such as the nature of the strategy of Buba and Kiki's community for dealing with indecision) in order to make sure that participants read the story carefully and understood it sufficiently well. Participants failing any of these questions were eliminated from the study.[15]

4.2.2. Hypotheses

We aimed to test two major hypotheses. First, as alluded to above, it is a straightforward prediction of the narrow view – more specifically: the version aiming to capture common intuitions regarding the concept of disagreement – that the relation between the individuals described in our story should not be understood as one of disagreement, at least not as long as participants maintain the assumption that Buba and Kiki did not change their view that the options are equally good. Call this H_1. Second, on the narrow view, those participants who nevertheless do *not* maintain this latter assumption (i.e., think the protagonist(s) *did* change beliefs) should be more likely to think that the protagonists disagree. Call this H_2.

To test H_1 and H_2, we looked at the data in the following way. First, to assess whether people tend to think that Buba and Kiki disagree without also thinking that they have conflicting beliefs, we only considered participants who said "no" to the view-change probe before counting their answers to the disagreement probe. Second, we compared these numbers with the group of people who said "yes" to the view-change probe in order to assess whether those who think Buba and Kiki's views differ are more likely to say they disagree. Furthermore, to address both H_1 and H_2 (and indeed all the hypotheses we tested), we also discounted the negative answers to the intention probe.[16]

There are two minor, additional hypotheses we tentatively attributed to defenders of the narrow view, both of which relate to the interpretation of our results in case they apparently go against the narrow view.

The first minor hypothesis – H_3—is that people may be liberal when applying the term "disagreement" to describe Buba and Kiki's situation even if they do not think that this is, strictly speaking, a case of disagreement, at least plausibly so if the disagreement probe appears first. Conversely, continues the hypothesis, if the view-change probe is presented first instead, this would induce a certain degree of "conceptual precision" and make participants more likely to judge there to be no disagreement ("strictly speaking"), since they have already had a chance to express the thought that Buba and Kiki do not have differing views.[17]

Finally, the other minor hypothesis – H_4—is grounded in the following thought: although participants may accept the idea that Buba and Kiki have no opposing beliefs about which option is better (after all, this is explicitly stated in the story), perhaps the only reason they state that Buba and Kiki disagree is that they wish to express the sense that there is a *non-disagreement-like clash* between them. Since the only way to do so is to give a positive answer to the disagreement question, such positive responses cannot be taken as evidence against the narrow view. In line with this thought, H_4 predicts that if participants have an opportunity to express their sense that there is a non-disagreement-like clash between Buba and Kiki before they answer the disagreement probe, the pressure to give a positive answer to the latter dissipates.

For this reason, we added a variation in our design in which the intention probe precedes the disagreement probe. According to H_4, the intention probe should function as a "valve" that allows expression to the perceived non-disagreement-like clash between Buba and Kiki.[18] As we explain below, the results show that all of these hypotheses are false.

4.2.3. *Results and discussion*

The majority of participants (214 of 357 or 60%) denied that, as a result of forming a new intention, Buba and Kiki changed their view that the options are equally good. Most participants also performed as expected on the intention probe, a clear majority (317 of 357 or 89%) stating that Buba and Kiki did end up having different intentions. By combining the "no" answers to the view-change probe and the "yes" answers to the intention probe, we obtained a sample of 203 participants to test H_1 (as well as H_3 and H_4). In this crucial sample, the overwhelming majority (173 of 203 or 85%) indicated that they believed Buba and Kiki disagree. Thus, contrary to H_1, even though Buba and Kiki's beliefs were not deemed to differ, there was a clear tendency to regard them as disagreeing with each other.[19]
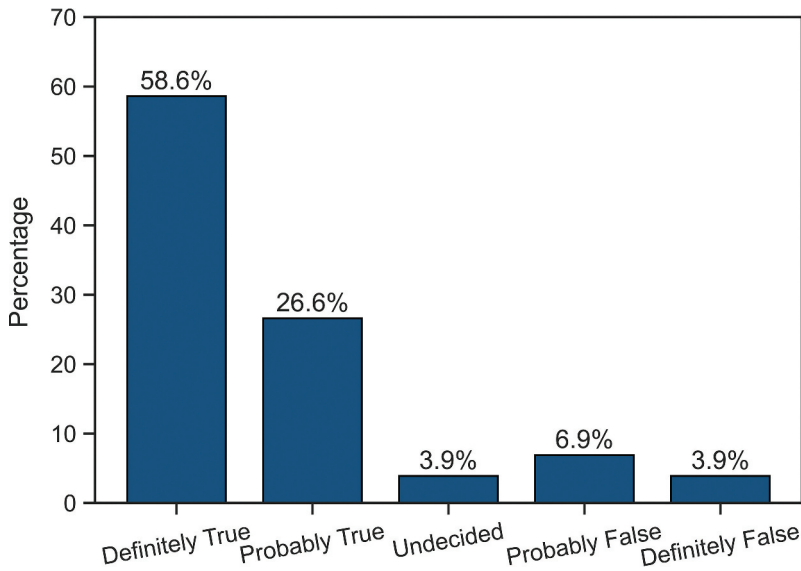
Figure 1. The distribution of responses to the disagreement probe among those participants who thought Buba and Kiki did not change their views about whether the options are equally good but ended up having different intentions (*N* = 203).

This strong tendency was also reflected in terms of confidence, which can be seen when the results are expressed in terms of the original 5-point scale: 119 of 203 (59%) believed that the statement "Buba and Kiki disagree about how to spend the evening together" was "definitely true", while 54 (26%) said it was "probably true". Only 22 (11%) answered negatively ("probably false" or "definitely false"), and there were only 8 (4%) participants who were undecided (Figure 1).

Regarding H$_2$, we found that those who saw Buba and Kiki's situation as involving a belief difference were not more likely to say they disagreed: 86 of 114 (75%) of those who said "yes" to the view-change probe and 173 of 203 (85%) of those who said "no" stated that Buba and Kiki disagreed (see Figure 2 below). In fact, as revealed by a chi-square test of independence, there was a slight but statistically significant tendency in the opposite direction: those whose responded with "yes" to the view-change probe (i.e., those who thought Buba and Kiki *did* change their minds) were significantly less likely to think Buba and Kiki disagreed; χ^2 (1, 317) = 4.67, *p* = .044, φ = .121.

We performed a logistic regression analysis to ascertain the effects of the relative ordering of the disagreement vs. view-change probes (DP-first vs. VP- first), the relative ordering of the intention probe vs. the other two probes (IP-first vs. IP-last), and the effect of the interaction of these two factors on the likelihood of a positive answer in response to the disagreement probe. Unlike the null model (see fn. 19), the logistic regression model
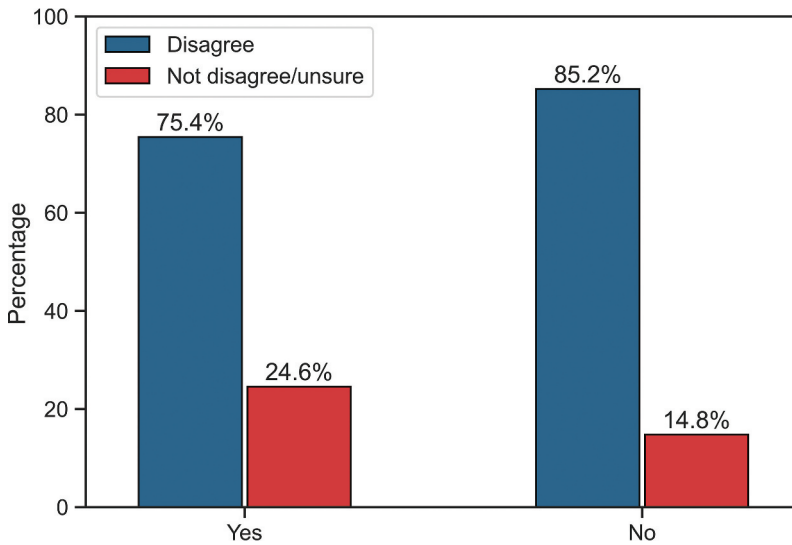
Figure 2. The distribution of responses to the disagreement probe as a function of the response to the view-change probe ("yes" or "no") among those participants, who thought Buba and Kiki ended up having different intentions ($N = 317$).

was not statistically significant, $\chi^2(3) = 1.42$, $p = .791$. Thus, contrary to H_3 and H_4, neither variation in terms of DP vs. VP, $p = .361$, OR $= .315$, 95% CI [0.03, 3.75], nor variation in terms of IP-first vs. IP-last, $p = .484$, OR $= .420$, 95% CI [0.04, 4.78] was a significant predictor of response to the disagreement probe, nor in fact was the interaction of these two predictor variables significant $p = .371$, OR $= 2.031$, 95% CI [0.42, 9.73] when controlling for the effect of all other variables.[20]

In conclusion, we have little or no reason to believe any of the four hypotheses discussed above. Most crucially, regarding H_1 and H_2, we have seen that a situation in which the relevant parties do not have any conflicting beliefs, such as Buba and Kiki's, may still be seen by most people as a case of disagreement. This strongly suggests the narrow view, according to which disagreements necessarily involve conflicting beliefs, is actually in tension with pre-theoretical intuition rather than supported by it.

These results seem promising, but a correspondent of ours objected that our probes – particularly the view-change probe – might miss the mark. One might willingly accept that most participants in the experiment know that Buba and Kiki retain their belief that the two options are equally *good*. However, our critic could insist that this leaves open the possibility of a disagreement in belief about what *should* be done. (Recall that the emphasis on "should" claims was central to Frances and Mattheson's understanding of action-disagreement discussed in §3.) Despite their equal evaluation of the options, perhaps the arbitrary picking process leads Buba to form the belief that they *should* go to a movie and Kiki to form the belief that they

should go to a restaurant instead. "We should do A, not B" seems to express belief in a contradictory proposition to "we should do B, not A", and this is where one could insist the real disagreement is identified by the participants.[21]

In light of this objection, we updated the vignette and probe and ran a modified version of our study to investigate whether participants attribute the disagreement to Buba and Kiki's conflicting beliefs about what should be done. We present this study and its results below.

4.3. Study 2

4.3.1. Method

To address the worry explained at the end of the previous section, we modified the design of the previous study in ways we explain below.

We recruited 681 native speakers of English via Amazon Mechanical Turk, of whom 216 were eliminated due to failed control questions. Our final sample thus consisted of 465 participants (260 male, 203 female, 2 other, M_{age} = 38.45, SD = 10.75). This study was also pre-registered (see osf. io/ectrw). We presented participants with the following modified vignette (modifications as compared to the previous study materials appear in italics):

> Buba and Kiki are friends. They live in a remote community in which being indecisive is seen as a terrible vice. Members of this community have developed strategies in order not to be indecisive. Whenever they face a choice between options that seem equally good to them, they pick one of the options arbitrarily and then they form an intention to act on that option—without changing their view that the options are equally good. *Since the options still seem equally good to them, they also don't think that one of the options should be pursued over the other. They merely intend to pursue the one they picked.*

> Buba and Kiki have decided to spend the evening together. The options available to them are to go to a movie or to go to a restaurant. Since, to both of them, these options seem like an equally good way to spend the evening together, each one of them resorts to the strategy for avoiding indecisiveness. Buba arbitrarily forms the intention to go to a movie and Kiki arbitrarily forms the intention to go to a restaurant.

> "I am only willing to go to a movie," says Buba. "I am only willing to go to a restaurant," says Kiki.

As previously, participants then were asked three main questions concerning (a) whether they thought Buba and Kiki disagreed (*disagreement probe*), (b) whether Buba and Kiki have changed their views about what should be done (*change of should probe*),[22] and (c) whether Buba and Kiki end up having different intentions at the end of the story (*intention probe*). (a) and

(c) were exactly as before, while (b) aimed addressing the issue that motivated this study (along with the added sentences in the vignette).[23]

Finally, as before, we included three control questions inquiring into some details of the story. Again, participants failing any of these questions were eliminated from the study (for more details, see osf.io/ectrw).

4.3.2. Hypotheses

According to the objection presented above, what is supposed to explain our previous results is that, although our participants did not think that Buba and Kiki had different beliefs regarding whether the options are equally good, they did think that Buba and Kiki have conflicting beliefs as to what should be done. *Ipso facto*, they do have relevant conflicting beliefs, which in turn can explain why they are judged to disagree. If so, we can revive the first two major hypotheses from the previous study. Call the new ones $H_1^\star$ and $H_2^\star$.

On $H_1^\star$, those participants who don't think Buba and Kiki change their views about what should be done are expected to reject the statement according to which Buba and Kiki disagree. On $H_2^\star$, those participants who think that Buba and Kiki changed their views as to what should be done should be more likely to think they disagree.

4.3.3. Results and discussion

A majority of participants (333 of 465 or 72%) denied that, as a result of forming a new intention, Buba and Kiki changed their view about what should be done. Most participants (412 of 465 or 89%) also agreed that Buba and Kiki ended up having different intentions.[24] By combining the "no" answers to the change of should probe and the "yes" answers to the intention probe, we obtained a sample of 300 participants to test $H_1^\star$. In this crucial sample, the majority of participants (245 of 300 or 82%) thought that Buba and Kiki disagreed.[25]

As Study 1, this tendency was also reflected in terms of confidence, which is apparent when the results are expressed on the original 5-point scale. 194 of 300 (65%) believed that the statement "Buba and Kiki disagree about how to spend the evening together" was "definitely true", while 51 (17%) said it was "probably true". Only 48 of 300 (16%) answered negatively (22: "probably false", and 26: "definitely false"), while 7 participants (2%) were undecided (Figure 3).

Regarding $H_2^\star$, we found that those who saw Buba and Kiki's situation as involving a belief difference were not more likely to say they disagreed: 61 of 112 (54%) of those who said "yes" and 245 of 300 (82%) of those who said "no" to the change of should probe stated that Buba and Kiki disagreed (see Figure 4 below). Much like in the previous study, a chi-square test of independence revealed that
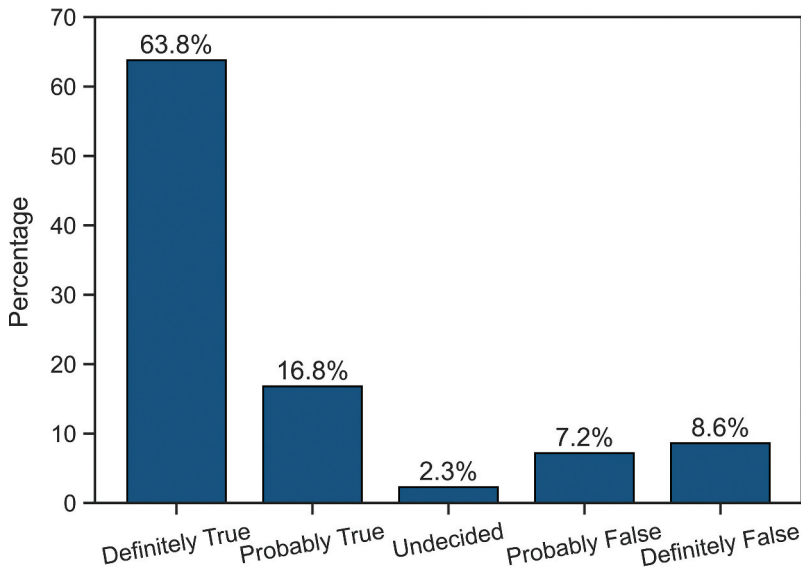
Figure 3. Responses to the disagreement probe among those participants who thought Buba and Kiki did not change their views about what should be done but ended up having different intentions (*N* = 300).
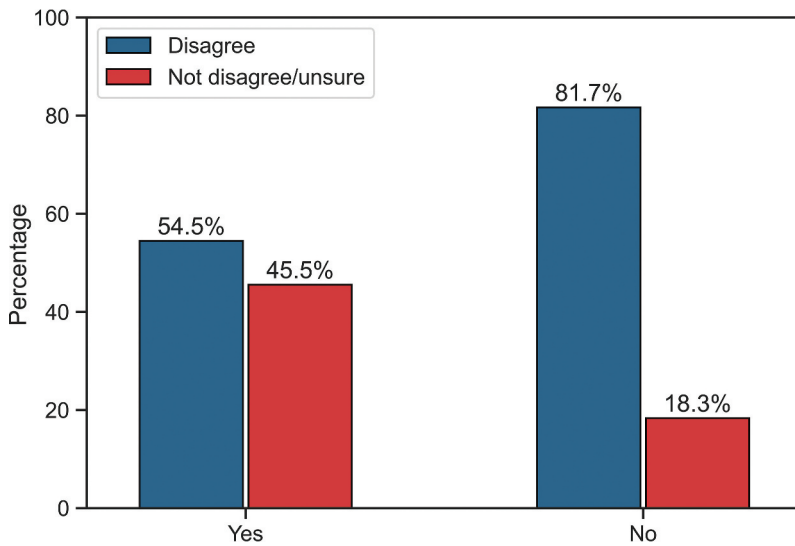

Figure 4. The distribution of responses to the disagreement probe as a function of the response to the view-change probe ("yes" or "no") among those participants, who thought Buba and Kiki ended up having different intentions (*N* = 317).

there was a statistically significant tendency in the opposite direction: those who responded with "yes" to the view-change probe were *less* likely to think Buba and Kiki disagreed; χ^2 (1, 412) = 31.58, *p* < .001, φ = .277.

Thinking that there is a difference between views (either about whether the options are equally good or about what should be done) does not make participants more likely to think the protagonists disagree. What about differences in terms of attributed intentions? If participants think Buba and Kiki changed their intentions (as speculated in the vignette), are they more or less likely to think Buba and Kiki disagreed? To explore this question, we included those who said "no" to the intention probe (as well as those who said "yes" to the change of should probe). Among those who said "no" to the intention probe, 28 of 53 (or 53%) believed that Buba and Kiki disagreed. Among those who said "yes" to the same question, this proportion increased to 306 of 412 (or 74%). As a chi-square test of independence revealed, this difference was significant; $\chi^2(1, 465) = 10.669$, $p = .001$, $\varphi = .151$.

In conclusion, again, we have provided evidence against both major hypotheses derived from the narrow view, namely, $H_1^\star$ and $H_2^\star$.[26] Taken together, the results presented in our two studies pose a robust challenge to the narrow view, according to which a clash of beliefs is necessary for a relation between two individuals to be seen as a disagreement.[27]

5. General discussion

We briefly address four questions:

(1) What further empirical research is called for?

Answer: We would like to see these results replicated. We would further like to see whether similar results will be found with other non-belief attitudes such as desires or emotions.[28] We would also like to see whether similar results will be found in other contexts. In particular, it would be interesting to check whether the results are stable across languages and geographical areas.

(2) Are a person's linguistic choices a good indication of her conceptual understanding?

Answer: We recognize that there often is a distinction drawn between one's way of applying terms in certain contexts and the extension of one's concepts. A person could, in theory, think that strictly speaking only conflicts in beliefs are disagreement but use the word "disagreement" to refer (loosely) to other conflicts as well. Why? Maybe because the distinction has no practical import in a given context, or maybe for some other reason. Doesn't, then, this recognition undermine our methodological assumption that language use is a good guide to the user's understanding of the associated concepts?

We think not!

The application of terms to situations is one of the best indicators of a person's conceptual understanding. (Indeed, this approach is standard in cognitive science, see e.g., Murphy, 2002 but for a voice of caution regarding this practice, see Malt et al., 2015). This indicator can and should be used to draw *prima facie* conclusions about the extension of a concept. One might pragmatically use a term in a way that does not reflect well the semantics of the relevant concept. If there are good reasons to suspect that some such pragmatic mechanism is at play in particular situations, these reasons should be presented and assessed. But the burden of proof is on those who doubt the validity of the inference from language use to conceptual understanding.

As a matter of fact, we have twice attempted to evaluate potential wedges that could tell against our results on the ground that ordinary usage of the term "disagreement" in the context of our study is less precise than the concept people implicitly hold. We did so in our assessment of H_3 and H_4. In both cases the potential wedges were not vindicated.

(3) What are the implications for noncognitivism in metaethics?

Answer: According to noncognitivists, normative beliefs are desire-like states. According to plan-expressivists (Ayars, 2021; Gibbard, 2003, 2012), they are strongly connected to intentions: they are decision-states that reliably result in intentions, they are intention-*like* states, or they are *identical* to intentions. "Murder is wrong" might express the plan to feel guilt and remorse upon committing murder and to blame people for murdering, for example. One disagreement objection to expressivism is that these kinds of intentional states don't really disagree with one another, so since there are clearly disagreements in normative belief, expressivism must be false. At first glance, our studies are helpful for plan-expressivists as they reveal common intuitions that there can be disagreement in intention. It may not be *entirely* helpful, though, because in our vignette Buba and Kiki are making plans *together*, and the more difficult case for the expressivist to capture is when people make different plans and are acting independently (Ayars, 2021, pp. 46–7; Bex-Priestley & Shemmer, 2017, pp. 192–3).

Conversely, is plan-expressivism a threat to our conclusion? Perhaps! It depends on how strong the view is. To see the threat, not only must we identify normative beliefs as intentions, but we must also identify the kind of intention that Buba and Kiki make as normative beliefs. This is unusual: expressivists typically don't want to identify *every* intention as a normative belief because sometimes we make plans without forming normative judgments about their superiority, maybe because the options seem equally good (Ayars, 2021, pp. 50–51), just like in Buba and Kiki's case, or we're under time pressure, or we just can't be bothered to weigh everything up. Yet if

Buba and Kiki's different intentions were in fact identical to normative beliefs, then their disagreement in intention *just is* disagreement in belief, and so the narrow view survives. Our conclusion, then, is technically on the proviso that very-strong-plan-expressivism is false. Nevertheless, even on this proviso our conclusion is interesting, because it implies at the least that there can be disagreements without any disagreement in *representational* belief.

(4) What is the import of our results for the debate about the nature of disagreement?

Answer: We observed earlier that there are no positive arguments for the view that disagreement can only be grounded in conflicts of beliefs. Our current results also suggest that defenders of belief-required disagreement cannot take their view as the default option supported by common sense intuition. These results must, of course, stand up to further experimental scrutiny. But if they do, then, the burden will be on defenders of the narrow view to make a theoretical case for the superiority of their position; a case that would be strong enough to override our findings that common sense actually supports the view that non-belief disagreement is genuine.

They may argue that views other than theirs are confused or non-unified or that, taken to their logical conclusions, they lead to absurd consequences; or that we should *choose to adopt* a belief-required view (to "ameliorate" the concept) for some reason, e.g., that it best serves some moral goal. However, our results suggest that, in the absence of such arguments, there is no reason to take the narrow view as the default position. Instead, theories that recognize disagreement beyond beliefs should be seen as doing better justice to pre-theoretical intuitions. Many such theories have been recently proposed (Beddor, 2019; Bex-Priestley & Shemmer, 2017; MacFarlane, 2014, ch.6; Ridge, 2013, 2014, ch.6; Shemmer & Bex-Priestley, 2021; Worsnip, 2019). Which of these theories is best is a further question to settle.

Notes

1. Notable exceptions are MacFarlane (2007), and Parfit (2011).
2. Here are some examples. Feldman and Warfield in the introduction to their edited collection about disagreement (Feldman & Warfield, 2010) list the major questions that arise in the philosophy of disagreement. Not only is the question of whether there could be disagreement without belief absent from the list, this option is not even mentioned as having a possible relevance to any of the questions that are listed. Larmore, in his essay "Pluralism and Reasonable Disagreement" (Larmore, 1994) doesn't consider even once the possibility that reasonable disagreement is ubiquitous because it is not belief-based. This example is particularly striking since non-belief disagreement is a prime candidate for explaining why reasonable political disagreement is widespread. Amy Gutmann and Dennis Thompson in their essay "Moral Disagreement in Democracy" (Gutmann & Thompson, 1995) are so convinced that

disagreement is a matter of divergent "opinions" that even when arguing against the view that disagreement always has its roots in conflicting interests, they merely interpret it as the view that conflicting interests explain the origins of divergent opinions. And the list goes on and on.

3. Later on, the article also recognizes disagreement in terms of differing levels of confidence about a certain belief.

4. While Frances and Mattheson wish to reduce such disagreements to disagreements in belief about what *should* be done, we do not think this is the most natural way to do it. For some attitudes it doesn't seem that there must be a particular one that people "should" have, and if there are disagreements in intention about supererogatory actions then framing these in "should"-terms looks implausible. It makes more sense to us to frame things in terms of which options and attitudes are *better* or *best*. For this reason, in our first study we assume that the reduction we are challenging is a reduction to beliefs about which option is better than the other. In the second study, we return to Frances and Mattheson's way of understanding the reduction in order to test that too.

5. These philosophers can be very roughly divided into two groups. Those who see belief and non-belief disagreement as fundamental but distinct relations, and those who see them as two cases of a single fundamental relation (Bex-Priestley & Shemmer, 2017, pp. 190, 196, 200).

6. For an overview and a critical reaction to that research programme see Hopster (2019), and Beebe (2021).

7. Beebe (2021, p. 82).

8. See our discussion on page 5 for a partial list of participants in this debate.

9. Note however, that this rejection is not a criticism of the empirical work on "folk objectivism" cited in the previous paragraph. Since researchers of folk objectivism are not interested in characterizing disagreement, as most often they merely think of the situations presented to respondents as a situation of disagreement without describing them in that way in the experimental prompts, and since on the rare occasions in which the term does appear in the experimental prompts it is only used to help participants identify the situation that needs to be assessed, there is nothing wrong with their stipulative usage.

10. Khoo and Knobe (2018, p. 111). On the same page the authors claim that this stipulative view of disagreement is the standard view. Unfortunately, their citations lack page numbers. We suspect that at least some of the authors they appeal to in order to support this claim would have disagreed.

11. Khoo and Knobe (2018, p. 118).

12. We emphasize again that this objection is not an objection to Khoo and Knobe's conclusion, but rather to the possibility of using their findings as evidence for the possibility of disagreement without belief. Incidentally, they do not say whether they understand the conflicting speech acts under discussion as expression of underlying beliefs or as expression of underlying non-belief attitudes.

13. The term "clash" does not appear in the vignettes or the questions presented to participants. We use it in our description of the studies to refer to cases where agents who are engaged in a common project have intentions to complete the project which cannot be simultaneously satisfied. A question we do not address in the current paper is what notion "clash" captures, and whether any single notion can capture the relation between all disagreeing attitudes. That question receives different answers by different philosophers of disagreement. Many of these answers, including our own,

can be found in the works mentioned on p. 5 in the introduction and in section 5.4 below.

14. The probes are as follows.

    (a) *Disagreement probe*:

    "To what extent do you believe the following statement is true?

    Buba and Kiki disagree about how to spend the evening together."

    The answers were given on a scale of 1 to 5 (1: "definitely true"; 2: "probably true"; 3: "undecided"; 4: "probably false"; 5: "definitely false"). For the purposes of the main analyses discussed below, we dichotomized these responses, taking 1 and 2 as positive answers to the disagreement probe, and taking 3–5 as negative answers.

    (b) *View-change probe*:

    "As a result of forming a new intention, did Buba and Kiki change their view that the options are equally good?" (Yes/No)

    (c) *Intention probe*:

    "Do Buba and Kiki end up having different intentions?" (Yes/No)

15. These control questions – as well as additional details of our design and a more complete presentation of the data – can be found at: osf.io/24qrc. Also note that in this paper, our focus is limited to a subset of the data. For example, beyond what is discussed below, we also collected open-ended qualitative data on participants' reasoning regarding their answer to the disagreement probe. Since these additional data do not affect the overall conclusion presented here, we have decided not to discuss them for reasons of space.

16. The intention probe may be seen as a manipulation check or an additional control question. After all, the whole point of the procedure of Buba and Kiki's community is to form a new intention.

17. Here, as in the case of H_1, we only looked at the negative answers to the view-change probe.

18. Thus, to address H_3 and H_4, the following order conditions were created (where "DP" = disagreement probe; "VP" = view-change probe; "IQ" = intention probe): (1) DP, VP, IP: (2) VP, DP, IP; (3) IP, DP, VP; (4) IP, VP, DP.

19. In a logistic regression analysis, which included the relative positions of the intention probe and view change probe (as well as the interaction of these two factors) as predictors, the null model was significant, meaning that irrespective of the predictor variables, the distribution of responses (85% vs. 15%) was significantly different from an even (or 50–50) distribution, $p < .001$ (see further below for more details of this analysis).

20. 88 of 104 (85%) said Buba and Kiki disagreed in the DP-first condition, compared to 85 of 99 (86%) in the VP-first condition. The numbers in relation to the IP-first and the IP-last conditions were the following: 94 of 109 (86%) and 79 of 94 (84%), respectively.

21. This assumes "should" statements express beliefs, and even then, it's not so straightforward because the contents of those beliefs might be context-dependent and compatible (see e.g., Finlay, 2017).

22. The new question read as follows: "Did Buba and Kiki change their views about what should be done during their evening together after changing their intentions? (Yes/No)".

23. Since no order effects were found in the first study, we merely included two order variations (disagreement probe first vs. change of should probe first) to perform

preliminary tests, thus, the design of this study is simpler than that of the previous one.

24. Preliminary tests indicated that the relative order of the disagreement probe and the change of should probe did not matter in terms of the responses to either of these probes. For the first, 155 of 215 (or 72%) responded with "yes" to the disagreement probe in the DP-first condition (i.e., where the disagreement probe was presented first relative to the change of should probe) compared to 151 of 197 (or 77%) in the DP-second condition. This difference was not significant; $\chi^2(1, 412) = 1.117$, $p = .29$. For the second, 161 of 215 (or 75%) responded with "no" to the change of should probe in the DP-first condition compared to 139 of 197 (or 71%) in the DP-second condition. Again, this difference was not significant; $\chi^2(1, 412) = 0.972$, $p = .32$.

25. This distribution differs significantly from what would be expected by chance, $\chi^2(1, 300) = 120.33$, $p < .001$.

26. In fact, the last two tests indicated that (1) participants who thought the protagonists changed their views on what should be done were *less* likely to think they disagreed, and (2) those who thought the protagonists had different intentions were *more* likely to think so. These results are consistent with those of Study 1. Whatever the explanation of these effects, the narrow view is clearly not well positioned to provide it.

27. A defender of the narrow view may try to save their position by making a distinction between the objective, evaluative "should" and the deliberative "should". They may insist that the participants in our studies believe the following: granted, Buba and Kiki both agree that their options are equally objectively evaluatively good, but they *disagree in belief* about what *deliberatively* should be done. However, assuming that these two distinct "should"s come apart in the required way, in Buba and Kiki's situation it seems natural to us to deny that either of the two options is what should be done in *either* sense of the word. The result of deliberation seems to us to result in neither answer, which is why Buba and Kiki resort to an arbitrary, rather than deliberative, method of picking what to do. The narrow view defender may suggest that once the arbitrary method has concluded, the deliberative "should" follows, based on a principle that one should do what one has arbitrarily picked in this situation. Yet it is now less clear that the respective beliefs of Buba and Kiki disagree: they both believe the principle of *one (deliberatively) should do what one has picked*, but they've both picked different things – so what do they disagree about?

　　As a final resort, the narrow view defender may say that the participants in our studies attribute unfairness to at least one of the disagreeing parties. Perhaps they think Buba believes "Kiki (deliberatively) should do what I have picked" and Kiki believes "I (deliberatively) should do what I have picked," and so even though Buba and Kiki agree that neither option objectively evaluatively should be done, they disagree in belief about what Kiki deliberatively should do. Our studies do not rule this out, and so we thank an anonymous reviewer for raising this possible escape route. We are skeptical that this is what the subjects believe – we think it is unlikely that the subjects attribute a belief to (say) Buba that Kiki (deliberatively) should go to a movie, for example – but we cannot be sure without further tests.

28. Finding, or failing to find, similar results with respect to other non-belief attitudes would enable us to delineate more exactly the scope of the concept of disagreement. It would not, however, change the fundamental import of the current findings. Even if intentions were the only non-belief attitudes which ground relations of disagreement, the traditional account would still need to be rejected.

ORCID

Miklós Kürthy http://orcid.org/0000-0002-7940-845X
Graham Bex-Priestley http://orcid.org/0000-0001-7731-5535
Yonatan Shemmer http://orcid.org/0000-0001-5842-8470

Data availability statement

The data that support the findings of this study are openly available in OSF at http://doi.org/10.17605/http://osf.io/wcpqt, reference number: http://osf.io/wcpqtosf.io/wcpqt.

References

Ayars, A. (2021). Deciding for others: An expressivist theory of normative judgment. *Philosophy and Phenomenological Research*, *105*(1), 42–61. https://doi.org/10.1111/phpr.12800

Beddor, B. (2019). Subjective disagreement. *Noûs*, *53*(4), 819–851. https://doi.org/10.1111/nous.12240

Beebe, J. R. (2021). The empirical case for folk indexical moral relativism. In T. Lombrozo, J. Knobe, & S. Nichols (Eds.), *Oxford studies in experimental philosophy* (Vol. 4, pp. 81–111). Oxford University Press.

Bex-Priestley, G., & Shemmer, Y. (2017). A normative theory of disagreement. *Journal of the American Philosophical Association*, *3*(2), 189–208. https://doi.org/10.1017/apa.2017.17

Björnsson, G., Finlay, S., Björnsson, G., & Finlay, S. (2010). Metaethical contextualism defended. *Ethics*, *121*(1), 7–36. https://doi.org/10.1086/656304

Blackburn, S. (1998). *Ruling passions*. Clarendon Press.

Dreier, J. (1999). Transforming expressivism. *Noûs*, *33*(4), 558–572. https://doi.org/10.1111/0029-4624.00193

Egan, A. (2014). There's something funny about comedy: A case study in faultless disagreement. *Erkenntnis*, *79*(S1), 73–100. https://doi.org/10.1007/s10670-013-9446-3

Feldman, R., & Warfield, T. A. (2010). Introduction. In R. Feldman & T. A. Warfield (Eds.), *Disagreement* (pp. 1–9). Oxford University Press.

Finlay, S. (2014). *Confusion of tongues: A theory of normative language*. Oxford University Press.

Finlay, S. (2017). Disagreement lost and found. In R. Shafer-Landau (Ed.), *Oxford studies in metaethics* (Vol. 12, pp. 187–205). Oxford University Press.

Frances, B., & Matheson, J. (2019). Disagreement. In E. N. Zalta (Ed.), *Stanford encyclopedia of philosophy (winter 2019 ed.)*. Stanford University. https://plato.stanford.edu/archives/win2019/entries/disagreement/

Gibbard, A. (2003). *Thinking how to live*. Harvard University Press.

Gibbard, A. (2012). *Meaning and normativity*. Oxford University Press.

Goodwin, G. P., & Darley, J. M. (2008). The psychology of meta-ethics: Exploring objectivism. *Cognition*, *106*(3), 1339–1366. https://doi.org/10.1016/j.cognition.2007.06.007

Goodwin, G. P., & Darley, J. M. (2010). The perceived objectivity of ethical beliefs: Psychological findings and implications for public policy. *Review of Philosophy and Psychology*, *1*(2), 161–188. https://doi.org/10.1007/s13164-009-0013-4

Goodwin, G. P., & Darley, J. M. (2012). Why are some moral beliefs perceived to be more objective than others? *Journal of Experimental Social Psychology*, *48*(1), 250–256. https://doi.org/10.1016/j.jesp.2011.08.006

Gutmann, A., & Thompson, D. (1995). Moral disagreement in a democracy. *Social Philosophy & Policy*, *12*(1), 87–110. https://doi.org/10.1017/S026505250000457X

Hopster, J. (2019). The meta-ethical significance of experiments about folk moral objectivism. *Philosophical Psychology*, *32*(6), 831–852. https://doi.org/10.1080/09515089.2019

Huvenes, T. (2012). Varieties of disagreement and predicates of taste. *Australasian Journal of Philosophy*, *90*(1), 167–181. https://doi.org/10.1080/00048402.2010.550305

Huvenes, T. (2014). Disagreement without error. *Erkenntnis*, *79*(S1), 143–154. https://doi.org/10.1007/s10670-013-9449-0

Khoo, J., & Knobe, J. (2018). Moral disagreement and moral semantics. *Noûs*, *52*(1), 109–143. https://doi.org/10.1111/nous.12151

Larmore, C. (1994). Pluralism and reasonable disagreement. *Social Philosophy & Policy*, *11*(1), 61–79. https://doi.org/10.1017/S0265052500004295

MacFarlane, J. (2007). Relativism and disagreement. *Philosophical Studies*, *132*(1), 17–31. https://doi.org/10.1007/s11098-006-9049-9

MacFarlane, J. (2014). *Assessment sensitivity: Relative truth and its applications*. Oxford University Press.

Malt, B. C., Gennari, S. P., Imai, M., Ameel, E., Saji, N., & Majid, A. (2015). Where are the concepts? What words can and can't reveal. In E. Margolis & S. Laurence (Eds.), *The conceptual mind: New directions in the study of concepts* (pp. 291–326). MIT Press.

Marques, T. (2015). Disagreeing in context. *Frontiers in Psychology*, *6*, 1–12. https://doi.org/10.3389/fpsyg.2015.00257

Marques, T., & García-Carpintero, M. (2014). Disagreement about taste: Commonality presuppositions and coordination. *Australasian Journal of Philosophy*, *92*(4), 701–723. https://doi.org/10.1080/00048402.2014.922592

McKenna, R. (2014). Shifting targets and disagreement. *Australasian Journal of Philosophy*, *92*(4), 725–742. https://doi.org/10.1080/00048402.2014.891145

Murphy, G. L. (2002). *The big book of concepts*. MIT Press.

Nichols, S. (2004). After objectivity: An empirical study of moral judgment. *Philosophical Psychology*, *17*(1), 5–28. https://doi.org/10.1080/0951508042000202354

Parfit, D. (2011). *On what matters, volume two*. Oxford University Press.

Priest, G. (2006). *In Contradiction* (2nd ed.). Oxford University Press.

Richard, M. (2015). What is disagreement? In *Truth and truth bearers: meaning in context* (Vol. 2, pp. 82–114). Oxford University Press.

Ridge, M. (2013). Disagreement 1. *Philosophy and Phenomenological Research*, *86*(1), 41–63. https://doi.org/10.1111/j.1933-1592.2011.00551.x

Ridge, M. (2014). *Impassioned Belief*. Oxford University Press.

Sarkissian, H., Parks, J., Tien, D., Wright, J. C., & Knobe, J. (2011). Folk moral relativism. *Mind & Language*, *26*(4), 428–505. https://doi.org/10.1111/j.1468-0017.2011.01428.x

Shemmer, Y., & Bex-Priestley, G. (2021). Disagreement without belief. *Metaphilosophy*, *52* (3–4), 494–507. https://doi.org/10.1111/meta.12489

Stevenson, C. L. (1944). *Ethics and language*. Yale University Press.

Sundell, T. (2010). Disagreements about taste. *Philosophical Studies*, *155*(2), 267–288. https://doi.org/10.1007/s11098-010-9572-6

Theriault, J., Waytz, A., Heiphetz, L., & Young, L. (2017). Examining overlap in behavioral and neural representations of morals, facts, and preferences. *Journal of Experimental Psychology General*, *146*(11), 1586–1605. https://doi.org/10.1037/xge0000350

Worsnip, A. (2019). Disagreement as interpersonal incoherence. *Research Philosophica*, *96* (2), 245–268. https://doi.org/10.11612/resphil.1771

Wright, J. C., Grandjean, P. T., & McWhite, C. B. (2013). The meta-ethical grounding of our moral beliefs: Evidence for metaethical pluralism. *Philosophical Psychology*, *26*(3), 336–361. https://doi.org/10.1080/09515089.2011.633751

Wright, J. C., McWhite, C. B., & Grandjean, P. T. (2014). The cognitive mechanisms of intolerance: Do our meta-ethical commitments matter? In T. Lombrozo, J. Knobe, & S. Nichols (Eds.), *Oxford studies in experimental philosophy* (Vol. 1, pp. 28–61). Oxford University Press.