Misunderstanding Metaethics:
Difficulties Measuring Folk Objectivism
and Relativism

– Lance S. Bush –
– David Moss –

Abstract: Recent research on the metaethical beliefs of ordinary people appears to show that they are *metaethical pluralists* that adopt different metaethical standards for different moral judgments. Yet the methods used to evaluate folk metaethical belief rely on the assumption that participants interpret what they are asked in metaethical terms. We argue that most participants do not interpret questions designed to elicit metaethical beliefs in metaethical terms, or at least not in the way researchers intend. As a result, existing methods are not reliable measures of metaethical belief. We end by discussing the implications of our account for the philosophical and practical implications of research on the psychology of metaethics.
Keywords: experimental philosophy, meta-ethics, metaphilosophy, moral cognition, moral objectivism, moral reasoning, moral relativism, validity

1. Introduction

Many studies purportedly provide evidence about the degree to which ordinary people endorse objectivism, relativism, and other metaethical beliefs[1], as well as factors that predict variance in metaethical belief[2] and how changes in metaethical beliefs influence attitudes and behavior.[3] However, no research paradigm exploring folk belief has provid-

Lance S. Bush
Cornell University
Department of Psychology
Uris Hall, G70
Ithaca, NY, 14853
lsb229@cornell.edu
David Moss
Canterbury Christ Church University
Old Sessions House
N Holmes Rd
Canterbury, CT1 1NX
david.moss@cantab.net

[1] E.g. Beebe (2015); Goodwin, Darley (2008; 2012); Nichols (2004); Sarkissian, Park, Tien et al. (2011); Wright, Grandjean, McWhite (2013).
[2] E.g. Beebe, Sackris (2016); Feltz, Cokely (2008); Goodwin, Darley (2008; 2012).
[3] E.g. Rai, Holyoak (2013); Yilmaz, Bahçekapili (2015); Young, Durwin (2013).

ed convincing evidence that participants understand what they are asked in metaethical terms, despite, as we will show, compelling reasons to believe that participants are not interpreting questions as intended. Crucially, we do not think that these problems are simply the result of avoidable errors in the design of these measures, but rather reflect inherent difficulties in specifying questions in a way that untrained respondents will interpret in line with the appropriate metaethical distinctions.

We explain why participants do not interpret questions about metaethics as intended in section §2. Metaethical distinctions are subtle and complex, and most people are unfamiliar with them. At the same time, there are multiple *alternative* interpretations of questions intended to be about metaethics which are more natural or intuitive than the interpretation intended by researchers. In section §3 we examine the methods used in existing research on folk metaethics. We show that they are riddled with potential confounds, due to them plausibly being interpreted as asking about non-metaethical questions or metaethical positions other than the ones researchers aim to study. We conclude by considering the implications of this proposal in section §4.

2. The Difficulties of Measuring Folk Metaethical Intuitions

There are a number of reasons to expect that few participants interpret questions about metaethics in line with the metaethical distinctions researchers are interested in measuring, including: *(i)* the relevant metaethical theories are complex and difficult to grasp, *(ii)* most people are unfamiliar with these distinctions prior to encountering them in studies, *(iii)* metaethical theories are generally *abstract* and distant from real world practical questions lay populations would be more familiar with and expect to be asked about, *(iv)* there are typically plausible *non-metaethical* interpretations of the questions posed to respondents, and participants may be more likely to interpret surveys as posing these more prosaic questions rather than abstruse metaethical questions. Taken together, we argue that these reasons result in responses being *less* likely to reflect the intended metaethical attitudes than non-metaethical (or unrelated metaethical) attitudes.

The claim that metaethical positions are complex and difficult to grasp should not be surprising to philosophers who have tried to teach students these theories. In our own pedagogical experience, which we think many philosophers will share, students often fail to properly understand these distinctions even after weeks of teaching. It might therefore seem that respondents drawn from the broader population, with no background in philosophy (and perhaps little interest in it), and usually a paragraph or less of text to introduce the theories, would be much less likely to accurately grasp them. A further reason to think that metaethical positions are complex and hard to understand comes from the observation that much philosophical debate involves professional philosophers arguing that other professional philosophers have failed to grasp a relevant philosophical distinction. If trained professionals struggle to grasp these concepts, it is unclear why we should suppose that untrained laypersons should readily do so. Despite this direct experience of others seeming to misunderstand metaethical theories, we suspect that researchers may systematically face a 'curse of knowledge', whereby

they are unable to appreciate the extent to which untrained individuals may struggle to correctly grasp these positions.[4]

To take the example of the most empirically studied metaethical issues, objectivism and relativism, it might seem like these can be characterized straightforwardly: 'are moral statements true only relative to particular standards or are they true objectively?'. Yet this apparent straightforwardness is deceptive. The literature has cashed this question out in at least two ways: "whether they [moral beliefs and standards] derive their truth (or warrant) independently of human minds (i.e., objectively), or whether instead, their truth is entirely mind-dependent or subjective" and "whether individuals treat their ethical beliefs as applying to all people, and all cultures."[5] Notably, these are non-equivalent: morality could be mind-dependent, but apply to all persons and cultures in the same way, or moral claims could apply differently across different persons or groups, but be mind-independent.[6] This introduces ample grounds for problematic confusion. For instance, in their research, Goodwin and Darley explicitly say that they are interested in mind-independence, *not* the scope of moral judgments.[7] However, if they or other researchers are interested in the *former* question, but respondents interpret the question as asking about the *latter*, or vice versa, then participant responses are simply not about the right topic.[8]

Yet the distinction drawn above is itself too vague to specify what kind of "objectivism" is being asked about, and teasing these possibilities apart requires the use of more technical philosophical terminology.[9] There are many different ways to be mind-dependent or mind-independent and not all of these are relevant to metaethical objectivism.[10] For example, beliefs about whether 'punching people is wrong' may depend on the mental fact that people tend to suffer pain when punched, or a judgement that an act was right or wrong may depend on the intentions of the actor, neither of which threatens mind-dependence in the relevant sense. Similarly, properties may be response-dependent (e.g. colors), even though statements about them are plausibly objectively true.[11] Determining whether morality is objective (in the sense of mind-independence) therefore may rely not only on grasping the right kind of mind-independence, but also on appropriately

---

[4] Camerer, Loewenstein, Weber (1989).

[5] Goodwin, Darley (2008): 1341. Other philosophers distinguish the question of mind-independence from the question of whether moral claims are indexical or not using the labels "objectivism" and "non-objectivism" and "relativism" and "absolutism" – Joyce (2015a). Note that while, according to this usage, objectivism and relativism are non-conflicting, according to the view of other philosophers "Metaethical moral relativist positions are typically contrasted with moral objectivism" – Gowans (2015).

[6] Joyce (2015a).

[7] Goodwin, Darley (2008): 1340. For an example of research that explores scope rather than mind-dependence, see Ayars, Nichols (2019): 2. They contrast relativism with universalism, defining the latter as "the view that there is a single true morality."

[8] This also leaves aside many other related distinctions drawn in the broader philosophical literature. For example, the distinction between descriptive, normative, and metaethical relativism – see Gowans (2015).

[9] Perhaps suggestive of the complexity and potential confusingness of this area, the philosophical literature has adopted a wide variety of incompatible and unclear terminologies – see Rysiew (2011).

[10] Joyce (2015b).

[11] Gert (2013).

distinguishing mental statements from the facts or properties they refer to and from the non-moral facts or properties which these supervene on or are constituted by.

Likewise, the question of the 'scope' of moral statements, whether they apply universally or whether they can be right or wrong for different groups and cultures, is ambiguous. Metaethical relativism concerns whether moral statements are 'indexical' i.e. relativized to the standards of some group or individual.[12] Yet moral statements or norms may apply to different groups or cultures in ways which do not entail metaethical relativism. For example, an act may be right in one context and wrong in another, due to a universal moral rule e.g. if it is universally morally wrong to needlessly insult people, it may be wrong to make a certain gesture in cultures where it is insulting to do so and morally acceptable to make the gesture where it is not insulting to do so. Moral statements may also be contextually-sensitive in various ways orthogonal to metaethical relativism.[13] For example, the meaning of the statement 'the hospital is not far' may depend on features of the context of assertion (e.g. whether we are walking or driving) despite facts about distance being objective. Likewise, the meaning of the statement 'it would be wrong for me to take that person's life' may also depend on features of the context of assertion (e.g. whether it would be an act of premeditated murder or an act of self-defense). In sum, correctly responding to questions about metaethical relativism requires interpreting them as semantic claims about the indexicality of moral statements, and not about other ways in which the truth of moral statements might vary across contexts, despite moral claims often appearing consistent with either interpretation.

We have argued that to respond to these questions appropriately, respondents need to be able to correctly grasp the correct metaethical interpretation out of a host of subtly different questions that might be being asked. How likely is it that respondents are able to do this? An additional barrier to participants' ability to do so is their lack of familiarity with the positions they are being asked about. Given that they will likely never have encountered the philosophical distinctions in question, it is unclear whether they should be expected to interpret the questions they are asked in studies to be about the particular metaethical question researchers have in mind, rather than some alternative interpretation consistent with the content of the question. In addition to their unfamiliarity, metaethical distinctions are also highly abstract and distant from the kinds of questions which respondents typically consider in everyday experience. Most ordinary moral judgments concern practical decisions about actual, concrete situations. Such situations invariably include a panoply of unique contextual details relevant to that situation. For instance, a typical moral question may be "Was it wrong for Alex to make that rude remark about our coworker last week at the work party?" Addressing this question would require knowing what was said, the context it was said in, and so on. It would be far less common for people to consider the moral status of rude remarks independent of any particular situation (a more abstract normative consideration), but there are even fewer circumstances in which people would be likely to consider *even more* abstract, metaethical questions, e.g. whether the truth status of moral sentences, as a category, are indexicalized.

---

[12] Joyce (2015b).
[13] Björnsson (2017): 4.

This is compounded by the availability of far more prosaic interpretations that ordinary people *would* be familiar with. For instance, consider a question that appears in metaethics research: participants are asked whether it would be wrong to fire a gun into a crowd of people on a busy city street.[14] Not surprisingly, most people think this is wrong. But they are then told that another participant disagreed with them, and are given the chance to affirm or deny that neither of them is mistaken (with an affirmative response taken to indicate rejection of objectivism). Participants are expected to interpret this as a situation in which the other person has different moral beliefs. But people may instead attribute their apparent moral disagreement to some cause other than a difference in moral beliefs. For instance, the other person may be thinking of a specific situation where using a gun would be justified, while the participant may be thinking of situations where it would not be. When asked to explain why they thought the other person disagreed with them, this is in fact what some participants did say. For instance, one participant stated that "The other person could be thinking about certain circumstances like the protection of others if there was a threat."[15] Of course, the frequency of these and other misinterpretations is an empirical question we are only beginning to address, but, as we argue below, available evidence supports our contention that such misinterpretations are more common than interpreting metaethical questions as intended.

3. Interpretative Difficulties in Metaethics Paradigms

3.1 The Disagreement Paradigm

The *disagreement paradigm* is the most common method for evaluating folk metaethical belief.[16] The precise framing of questions and response options vary, but all versions present participants with a moral disagreement between two people, then ask some variation of whether both people can be correct or if at least one person must be incorrect. For instance, Beebe and Sackris[17] ask participants to rate how much they agree with a series of moral statements (e.g. "Hitting someone just because you feel like it is wrong"), then ask:

> "If someone disagrees with you about whether [claim], is it possible for both of you to be correct or must one of you be mistaken?"
> □ *It is possible for both of you to be correct*
> □ *At least one of you must be mistaken*

In a seminal paper employing the disagreement paradigm, Goodwin and Darley argue that answers to this question indicate whether individuals believe moral standards are mind-independent or mind-dependent, which they refer to as "objectivism" and

---

[14] Goodwin, Darley (2008).
[15] Ibidem.
[16] E.g. Ayars, Nichols (2019); Beebe, Sackris (2016); Goodwin, Darley (2008; 2012); Sarkissian, Park, Tien et al. (2011); Wright, Grandjean, McWhite (2013).
[17] Beebe, Sackris (2016).

"subjectivism" respectively. They explicitly distinguish this from the question of whether moral standards apply to everyone or only to some individuals or groups, which they refer to as "universalism" and "relativism" respectively.[18] However, most later papers do not explicitly draw this distinction, and while some refer to mind-independence[19], others refer to testing whether moral claims are "relative" to particular standards or whether there is a single, "objective" moral standard.[20] Moreover, the terminology used across studies in this area is inconsistent, with some studies contrasting "objectivism" with "non-objectivism"[21] or "subjectivism,"[22] while other contrast "absolutism" with "relativism,"[23] and still others refer to "universalism" and "relativism,"[24] even when referring to earlier studies that used different terms, essentially treating the terms as interchangeable. As such, it is unclear where researchers are merely using different terminology to refer to the same concepts, where they might be testing substantively different hypotheses, and where they might be conceptually confused.[25] For simplicity, in the discussion below, we will simply refer to a contrast between "objectivism" and "relativism." This does not change the structure of our argument, because we claim that the disagreement paradigm is neither an adequate measure of whether individuals believe moral judgments are mind-independent nor whether they believe they are indexical.

For researchers using the disagreement paradigm, a judgment that 'It is possible for both of you to be correct' is interpreted as evidence of relativism, while the judgment that 'At least one of you must be mistaken' is interpreted as evidence of objectivism. These interpretations are based on the assumption that, to a relativist, seemingly conflicting claims such as 'abortion is wrong' and 'abortion is not wrong' can both be true as long as they are made relative to different moral standards (such as the subjective beliefs of different individuals or the standards of different cultures). In contrast, it is assumed that objectivists would believe that there is only one nonrelative standard of moral truth, so these statements represent mutually incompatible truth claims that could not be true at the same time.

---

[18] Goodwin, Darley (2008): 1341.
[19] Beebe, Qiaoan, Wysocki et al. (2015); Goodwin, Darley (2008; 2010); Pölzler, Wright (2020; 2019).
[20] Heiphetz, Young (2017); Sarkissian, Park, Tien et al. (2011); Sarkissian, Phelan (2019); Wright, Grandjean, McWhite (2013); Viciana, Hannikainen, Torres (2019).
[21] Feltz, Cokely (2008); Nichols (2004); Wright, McWhite, Grandjean (2014).
[22] Goodwin, Darley (2008); Fisher, Knobe, Strickland et al. (2017).
[23] Rai, Holyoak (2013).
[24] Ayars, Nichols (2019); Rose, Nichols (2019).
[25] A possible explanation for authors treating the terms as essentially interchangeable is that they may view responses to the disagreement paradigm as serving as proxies for judgments about *both* mind-independence and indexicality. A possible rationale for this view could be thinking that, despite the distinctions being conceptually orthogonal, in practice, lay respondents are likely to believe *either* that there is a single, mind-independent moral reality, to which moral statements non-indexically refer (objectivist absolutism), *or* that moral standards are mind-dependent and so moral statements indexically refer to these mind-dependent standards (non-objectivist relativism). Thus they might think that, although it is conceptually possible that respondents believe that moral statements are mind-dependent, but not relativist, or that they are mind-independent, but not objectivist, it is unlikely that respondents actually would hold these combinations of positions. We do not have space to adjudicate whether this is a reasonable presupposition, and whether it is actually true is clearly an empirical question. Nevertheless, this clearly poses a further potential complication for research in this area.

Studies employing the disagreement paradigm consistently find that most participants judge some moral disagreements as objective and other as relative.[26] For example, many participants judge that if two people disagree about whether abortion is permissible, they can both be correct (a relativist response) but also judge that if two people disagree about whether racial discrimination is wrong, at least one of them must be incorrect (an objectivist response). A handful of participants express uniformly objectivist or relativist responses, but they always represent a small minority of participants. This is taken to suggest that most ordinary people are metaethical *pluralists* that regard some moral issues as objective and others as relative.[27]

However, critics have identified numerous methodological concerns with the disagreement paradigm.[28] One line of criticism highlights the disagreement paradigm's inability to identify which metaethical beliefs people hold even if they understand what they are being asked. These criticisms have been developed at length elsewhere[29], and focus on problems in study design that may be avoidable. Yet these shortcomings are secondary to a more serious problem that is much harder to avoid: there is little reason to believe participants interpret the disagreement paradigm in metaethical terms in the first place. On the contrary, when participants were asked to explain their answers, they frequently interpreted the source of disagreement in ways inconsistent with a metaethical interpretation.

One way participants can misinterpret the disagreement paradigm is if they do not regard the source of the disagreement to be a difference in moral beliefs, but instead imagine that the source of the disagreement is due to some other cause, e.g. that one of the people misunderstood the question, meant something different than the other person, were referring to a different situation, etc.[30] To circumvent this, Goodwin and Darley asked people to explain what they thought the source of disagreement was and excluded anyone who did not interpret the disagreement "in a bona fide way."[31] They found that only "seven out of a total of 102 responses were excluded on these grounds" (i.e. 6.9%).[32] We decided to analyze their responses ourselves. Unfortunately, we did not reach the same conclusions. We coded all responses as either *correct*, *incorrect*, or *indeterminate* (i.e., interpretations that did not clearly demonstrate a correct or incorrect interpretation).

To interpret the source of disagreement correctly, participants simply had to attribute the disagreement to a difference in moral attitudes. Some did this, e.g. "the other person may have different values than I do." But such responses accounted for only about 41% of explanations, while most appeared to interpret the source of disagreement in an unintended way (44%) or at best did not appear to clearly demonstrate the intended interpretation (15%).[33] When participants misinterpreted the source of disagreement, they

---

[26] Wright, Grandjean, McWhite (2013).
[27] Wright (2018).
[28] Beebe (2015); Moss (2017); Pölzler (2018).
[29] E.g. Pölzler (2018).
[30] Goodwin, Darley (2008).
[31] Ibidem: 1348.
[32] Ibidem: 1348.
[33] All percentages reported in this paper are the first author's. We coded 100 of the 102 responses, since 2 were not available in our dataset.

often speculated that the person who disagreed with them must have been thinking of specific instances in which an otherwise immoral act would be permissible, e.g. one person suggested that if someone disagreed with them about whether it was wrong to rob a bank, this could be because "This person probably has specific details of such a happening where there were extreme circumstances that lead him/her to believe robbing a bank was not morally bad." However, in many cases we could not determine whether they understood the disagreement as intended. For instance, participants frequently condemned the other person as immoral, e.g. "Because the other person has clearly misplaced his or her conscience," or offered an etiological explanation, e.g. "Education, cultural background, financial situation." These responses are consistent with a correct interpretation, but neither clearly expresses one. For instance, if the other person misplaced their conscience, this could mean that they hold immoral beliefs, but it could also mean that although they share the same moral beliefs, they failed to judge in accordance with them on this occasion.

Beebe's findings point to another way people misinterpret the disagreement paradigm.[34] Beebe compared responses to moral disagreements to disagreements about empirical claims, such as the age of the earth. Most people recognize that some empirical claims have a single established and verifiable answer, a finding reflected in the high "objectivism" rates for these disagreements. However, Beebe also included empirical claims that lack any ready means of verification; or, as he puts it, are "not only unknown but practically unknowable."[35] One example involves conflicting claims about whether Julius Caesar drank wine on his 21st birthday.[36] If participants interpreted this question as intended, they would have to consider whether two logically incompatible historical claims could both be true *at the same time*. It is possible that some participants endorse a radical form of relativism about historical events. Nichols[37] found that some people appear to endorse relativism of this kind, although such people seem to be rare. Yet Beebe found that a whopping 45% of US participants judged that both people could be correct if they disagreed about whether Caesar drank wine on his 21st birthday, while Beebe et al. found a similar pattern in China (53%), Poland (48%), and Ecuador (53%).[38] Beebe also found high levels of "relativist" responses for whether exercise helps people lose weight (75%), whether humans are the primary cause of global warming (61%), and whether there are an even number of stars in the universe (44%), but not for claims that have well-established answers, e.g. whether New York City is further north than Los Angeles (8%), findings which were generally replicated across cultures. Should we really believe that nearly half of the world's population believes there is *no objective truth* about whether an historical event did or did not happen? If the disagreement paradigm is a valid measure of belief in objectivism and relativism, either this improbable finding is true, or there are at least some circumstances in which the disagreement paradigm cannot reliably distinguish objectivism from relativism.

---

[34] Beebe (2015).
[35] Ibidem: 393. See also Beebe, Sackris (2016).
[36] Later studies used similar scenarios (e.g. whether Caesar ate soup) or were adapted to be more culturally appropriate (e.g. referring to Confucius instead of Caesar).
[37] Nichols (2004).
[38] Beebe, Qiaoan, Wysocki et al. (2015).

One possibility is that participants interpret the disagreement paradigm, at least in some cases, to be asking them whether *(i) one of the two people is correct, so the other must be incorrect*, or *(ii) it is uncertain which of the two people is correct, so either one might be correct (but not both)*. If you are uncertain about which side of a dispute is correct, *(ii)* is an entirely reasonable response, but it has nothing to do with metaethics. The possibility that some people interpret it this way is bolstered by Wainryb et al.'s findings.[39] When asked to explain why they judged that two people could be correct when they disagreed about ambiguous factual issues that they could not know the answer to, e.g. why a dog didn't eat its food, 28% of 5-year-olds, 41% of 7-year-olds, and 66% of 9-year-olds attributed their answer to what Wainryb et al. described as 'uncertainty.' For example, one child said that "They can both be right because there's no way to know for sure, maybe the dog is hungry and maybe it doesn't like the food."[40] This suggests that, when participants are uncertain about which person is correct, they often interpret the disagreement paradigm to be asking an epistemic question, since their judgment that both can be correct reflects *(ii)* rather than the belief that the claim could be true relative to one standard and false relative to another.

This provides a plausible alternative explanation for the otherwise puzzling correlation between "objectivist" responses and perceived consensus.[41] That is, when there is widespread agreement that something is morally wrong, e.g. murder, most participants judge that only one person can be correct. But when the issue is highly contentious, e.g. euthanasia, people are more disposed to judge that people who disagree can both be correct. Although alternative explanations exist, such as Ayars and Nichols's[42] suggestion that people use consensus as evidence of objectivity, our account is more consistent with the explanations people give for their answers and offers a more plausible explanation for the high rate of "relativist" responses for uncertain nonmoral empirical claims. These are just two of the possible ways people may have misinterpreted the disagreement paradigm, but there are enough other ways the paradigm can be misinterpreted[43] paradigm that the cumulative effect of so many potential interpretative pitfalls suggests that more people misinterpret the disagreement paradigm than interpret it as intended. Fortunately, there are other methods for evaluating folk metaethical beliefs. But, as we will see, they do not fare much better.

3.2. Surveys

Yilmaz and Bahçekapili[44] use an adapted version of Forsyth's *Ethics Position Questionnaire* (EPQ) to evaluate folk metaethical belief.[45] Participants are given a list of statements ostensibly expressing objectivism or relativism and are asked to express how strongly

---

[39] Wainryb, Shaw, Langley et al. (2004).
[40] Ibidem: 692.
[41] Goodwin, Darley (2012).
[42] Ayars, Nichols (2019).
[43] See e.g. Pölzler (2018).
[44] Yilmaz, Bahçekapili (2015; 2018).
[45] Forsyth (1980).

they agree or disagree with each statement. Unfortunately, all of the items are ambiguous between metaethical and non-metaethical interpretations, such as descriptive claims about the degree of moral diversity in the world, or normative claims about the consequences of moral disagreement. For instance, one item states that "*What is moral varies on the basis of context and society.*" This could be a metaethical statement asserting that the truth values of moral statements vary relative to different contexts and societies, but it is also plausibly interpreted as an empirical observation that different societies have different moral beliefs (i.e. *descriptive relativism*). All items on the scale similarly conflate metaethical claims with non-metaethical claims. See **Table 1** for a summary of these conflations. Given that no item on this scale is face valid, it is not an appropriate tool for measuring metaethical belief.

Table 1: New Meta-Ethics Questionnaire[46]

Items	Conflations/Ambiguities with items
1. What is moral varies on the basis of context and society.	Conflates descriptive relativism with metaethical relativism. Conflates relativism with contextualism.
2. Moral standards are personal, therefore something morally acceptable to one person might be immoral for another person.	Conflates descriptive relativism with metaethical relativism.
3. Since moral rules are not absolute, no definite judgments about them are possible.	Conflates exceptionless rules, insensitivity to context, and/or universality with objectivism with the use of "absolute". Conflates epistemic and metaphysical interpretations with use of "definite."
4. Different cultures adopt different values and no moral law is right or wrong in an absolute sense.	Conflates descriptive relativism with metaethical relativism. Forces participant to agree/disagree with a compound statement. Conflates exceptionless rules, insensitivity to context, and/or universality with objectivism with use of "absolute".
5. We can agree on 'what is moral for everyone' because what is moral and immoral is self-evident.	Conflates epistemic and metaphysical interpretations with use of "self-evident." Conflates universality with objectivism.
6. If morality were to differ from person to person, it would be impossible for people to live together.	This is a question about the consequences of descriptive relativism. It is not related to metaethical objectivism/relativism.
7. Since the moral laws I believe in are universally true, they can be applied to everyone in the world regardless of culture, race or religion.	Conflates universalism with objectivism. Implies imposition of one's values on other people/cultures, which entangles normative considerations with metaethical ones.
8. If a moral law is right and good for others, it is also right and good for us.	Conflates normative and metaethical questions. Conflates universalism and objectivism. Implies imposition of one's values on other people/cultures, which entangles normative considerations with metaethical ones.

---

[46] Yilmaz, Bahçekapili (2015).

Yilmaz and Bahçekapili[47] recognized the poor face validity of these items and developed a new set of items in addition to modifying their original scale in ways that improve its face validity.[48] More recently, Collier-Spruel et al.[49] also developed an improved scale for evaluating relativism, while Zijlstra[50] created a more comprehensive scale that assesses other metaethical beliefs alongside relativism. Although all of these efforts circumvent the EPQ's poor face validity to varying degrees, when we asked respondents to explain why they agreed or disagreed with statements related to objectivism and relativism from each of these scales, many still consistently failed to interpret these items in metaethical terms.

For instance, Yilmaz and Bahçekapili[51] created a set of three items that are intended to be more face valid than items on the EPQ. However, when asked to explain why they agreed or disagreed with these statements, we found that few participants gave reasons that would indicate a metaethical interpretation. For instance, participants often interpreted *"There are no ethical principles that are so important that they should be a part of any code of ethics"* as a statement about whether our moral standards should be explicitly codified in some way, e.g. "I think that in order to ensure human rights, we need to make sure that they are in writing so that they are harder to violate," yet few interpreted the statement in metaethical terms (i.e. as a statement about whether moral standards are true in some mind-independent way that transcends culture and subjective belief).

Similarly, Zijlstra's relativist statements were rarely interpreted in metaethical terms, but were instead interpreted in a host of unrelated ways, e.g. much like the disagreement paradigm, participants often interpreted agreement with these items to reflect sensitivity to situational factors rather than relativism. For example, in explaining their agreement with the statement, 'There is not one but many different answers to the question of what is morally right and wrong and these can be equally correct' one participant stated that "Sometimes there are extenuating circumstances that influence actions. Until everything is known, it's not possible to say that there is only one morally correct course of action." This is an epistemic interpretation of the statement, *not* a metaethical one. Although rates of misinterpretation varied across these items, they are consistently high enough to warrant considerable skepticism about the validity of more recent scales that measure level of agreement with metaethical statements.

3.3. Direct Questions

Recognizing that participants may have misinterpreted the disagreement paradigm, Fisher et al. adopted a more direct approach.[52] They simply asked participants to consider the following question:

---

[47] Yilmaz, Bahçekapili (2015).
[48] Yilmaz, Bahçekapili (2018).
[49] Collier-Spruel, Hawkins, Jayawickreme et al. (2019).
[50] Zijlstra (2019).
[51] Yilmaz, Bahçekapili (2015).
[52] Fisher, Knobe, Strickland et al. (2017).

> *"'Should [action] be allowed?' Please tell us whether you think there is an objectively true answer to this question."*
> 1 (*Definitely no objective truth*) — 7 (*Definitely objective truth*)

Does directly asking people whether they think there is an 'objectively true' answer provide a valid measure of metaethical belief? Only if people understand the term 'objective' in line with what Fisher et al. mean by 'objective'. Fisher et al. state that objectively correct answers are ones that are "established by facts independent of any particular person's judgment" (p. 3). Like Goodwin and Darley, this suggests that they identify objectivism with mind-independence.

We decided to test whether ordinary people use the term 'objective' in this way. We recruited 57 participants through Amazon's Mechanical Turk (MTurk). Participation was limited to adults living in the US. No additional demographic data was collected. Participants were asked "*In your own words, what does it mean to say that moral truth is objective?*" A text box was available below this question where participants could provide a written response. No additional instructions were given.

Our goal in coding responses was to estimate the proportion of participants that interpreted 'objective' to refer to mind-independence and the proportion that interpreted 'objective' in some other, unintended way. However, many items were ambiguous or difficult to interpret. Since it would not be appropriate to code items that could not be confidently categorized as intended or unintended interpretations, these responses were categorized separately. All items were initially coded by the first author. 7 responses (12.3%) were judged as intended interpretations while 41 (71.9%) were judged to be unintended interpretations and another 9 (15.8%) were indeterminate. One example of an item initially coded as a clearly intended interpretation was, "it means it's a fact, not an opinion." However, it's questionable whether even this is a proper expression of objectivism; after all, relativist moral claims are also "facts."[53]

After all items were independently coded by the second author, we attempted to resolve disagreements via discussion where possible, though in some cases we were unable to reach consensus. This resulted in a reduction in the number of items coded as intended interpretations by the first author, with the first author coding 3 responses (5.3%) as intended interpretations, 45 responses (78.9%) as unintended, and 9 as indeterminate (15.8%), and the second author judging 0 responses (0%) as intended interpretations, 44 responses as unintended interpretations (77.2%), and 13 as indeterminate (22.8%).[54]

Even the few remaining items coded by the first author as intended interpretations may be insufficiently clear. For instance, one response stated that, "There are some absolutes, definite delineation between right and wrong, not subject to subjec-

---

[53] Even interpretations that seem unambiguously correct because they use explicitly metaethical language may not be accurate. For example, one respondent stated that "It means that moral truth is not subjective." Interpreting this as an accurate response relies on the assumption that the participant's notion of 'subjective' corresponds to some version of metaethical relativism. Yet other data we have collected (not reported here) suggests most participants also struggle to explain what it means to say that moral truth is relative.

[54] We achieved 91.2% agreement between coders and moderate interrater reliability, Cohen's Kappa = 0.764 (McHugh. 2012).

tive opinion." Given the latter remark that objective moral truths are "not subject to subjective opinion," this response could indicate an understanding of objective truth as mind-independent. However, the rest of this response is consistent with alternative interpretations, e.g. the notion that there is a "definite delineation between right and wrong" associates objectivism with the view that moral issues can be judged as either categorically right or wrong, which would be an unintended interpretation. Thus, even when being as charitable as we felt reasonable, it is still unclear whether *any* respondents associated objectivism with mind-independence.

While some items were difficult to interpret, it is nevertheless clear that the majority of participants did not interpret objectivism to refer to mind-independence. In fact, 23 (40.4%) respondents mistook objectivism for relativism or variation in moral beliefs. For example, one respondent stated that "Moral truth is objective because everyone has different views on what is moral or not. What might be immoral to one person might be perfectly moral to another person." In other words, around 40% of participants interpreted 'objectivism' to mean something more closely resembling the *opposite* of what Fisher et al. intended. Others suggested that moral truth is objective when judgment is unbiased, e.g. "It means what you believe is your truth and when you use your moral judgement it is unbiased." Notably, this *is* a common meaning of 'objective,' and such respondents have offered an entirely appropriate explanation of what 'objective' means; it just isn't the one that researchers asking the question had in mind. This exemplifies how misinterpretations are not merely the result of conceptual or linguistic failures on the part of participants, but are instead due in part to ambiguity and underspecificity in the questions themselves. Given these findings, it seems unlikely that Fisher et al.'s direct question provides a valid measure of metaethical belief.

4. Conclusion

We have argued that there are good reasons to believe folk responses frequently fail to reflect their stances towards the metaethical questions which researchers have aimed to investigate. What follows from this? One implication is that we cannot infer anything about folk metaethical psychology if our instruments are not measuring folk metaethical attitudes. Researchers therefore need to turn to the more fundamental task of conducting research to confirm participant understanding and demonstrate the validity of their instruments before they can confidently draw conclusions about the nature of folk metaethical thought. This is especially so for researchers who interpret their findings as support for the conclusion that folk metaethical views are irrational[55] or as evidence of metaethical pluralism.[56] While we are sympathetic to these accounts, if comprehension rates are as low as we propose, then the evidential grounds motivating irrationalism and pluralism are simply vitiated, since people's responses don't express metaethical attitudes in the first place. Unless this possibility can be conclusively dismissed, our alternative explanation that folk responses only appear varied or irrational because they are not appropriately understanding the questions plausibly stands as a *default* interpretation.

---

[55] Colebrook (under review).
[56] Wright, Grandjean, McWhite (2013).

Some researchers have drawn succor from the fact that replications of experimental philosophy studies have found relatively high replication rates.[57] Yet this sense of security is misplaced. Direct replications of previously reported effects do little, if anything, to assuage the kinds of concerns about validity we have raised here. An effect found using a measure that is systematically confounded or entirely misunderstood may be replicated perfectly, but will still be entirely invalid. To put it another way, one can repeatedly find the same effect without coming any closer to knowing what it *means*.

Fortunately, this state of affairs offers an opportunity for methodological innovation and for improving the validity and reliability of our instruments. One promising avenue is to adopt qualitative methodologies, such as semi-structured interviews[58] and more detailed analysis of open-comment data. These methods can serve not merely to confirm or disconfirm comprehension (as per cognitive interviewing[59]), but also as promising methods for investigating folk metaethics in their own right.

Another promising innovation is making greater use of 'training paradigms.' These studies aim to train philosophically naïve respondents in the distinctions that researchers are investigating so that (in theory) they are able to indicate their attitudes towards these questions.[60] However, this approach has a heavy theoretical cost, as instructing 'folk' respondents in philosophical concepts arguably moves us beyond what the folk *do* think about morality to what they *would* think if they began to be trained in philosophy. Furthermore, it is open to question whether philosophically naïve respondents can be induced to understand metaethical concepts within the space of a short survey. More research may be needed to confirm this.

Finally, even if it transpires that no method is successful in eliciting and measuring folk stances towards metaethical debates such as objectivism vs. relativism, there may still be valuable research to be done related to these areas. Even if philosophically untrained respondents simply cannot be brought to understand the relevant metaethical positions appropriately, for example, because they cannot distinguish the question of whether something is objectively true, from the question of whether it is truth-apt or universally true, they may still be able to identify a vague cluster of such positions, and avow that certain moral issues are more like straightforward matters of fact (e.g. whether the earth is flat) and others are more like matters of taste (e.g. whether chocolate is tasty). If so, then further research may identify important patterns in moral thinking, and metaethicists can plausibly make important contributions to these investigations.

Acknowledgements

---

[57] Cova, Strickland, Abatista et al. (2018).
[58] Moss (2017).
[59] Collins (2003); Andow (2016).
[60] Pölzler (2018); Wright (2018).

References

Andow J. (2016), "Qualitative Tools and Experimental Philosophy," *Philosophical Psychology*, 29 (8): 1128–1141.

Ayars A., Nichols S., (2019), "Rational Learners and Metaethics: Universalism, Relativism, and Evidence from Consensus," *Mind & Language* 35 (1): 67–89.

Beebe J.R., Sackris D. (2016), "Moral Objectivism Across the Lifespan," *Philosophical Psychology* 29 (6): 912–929.

Beebe J.R. (2015), "The Empirical Study of Folk Metaethics," *Etyka* 50: 11–28.

Beebe J.R., Qiaoan R., Wysocki T. et al. (2015), "Moral Objectivism in Cross-Cultural Perspective," *Journal of Cognition and Culture* 15 (3–4): 386–401.

Björnsson G. (2017), "Contextualism in Ethics", [in:] *International Encyclopedia of Ethics*, H. LaFollette (ed.), published online: 21 June 2017, doi: 10.1002/9781444367072. wbiee760.pub2.

Camerer C., Loewenstein G., Weber M. (1989), "The Curse of Knowledge in Economic Settings: An Experimental Analysis," *Journal of Political Economy* 97 (5): 1232–1254.

Colebrook R. (under review), "Reconceptualizing Folk Metaethics," URL== https://www. academia.edu/36014543/The_Irrationality_of_Folk_Metaethics [Accessed 8.5.2020].

Collier-Spruel L., Hawkins A., Jayawickreme E. et al. (2019), "Relativism or Tolerance? Defining, Assessing, Connecting, and Distinguishing Two Moral Personality Features with Prominent Roles in Modern Societies," *Journal of Personality* 87 (6): 1170–1188.

Collins D. (2003), "Pretesting Survey Instruments: An Overview of Cognitive Methods," *Quality of Life Research* 12 (3): 229–238.

Cova F., Strickland B., Abatista A. et al. (2018), "Estimating the Reproducibility of Experimental Philosophy," *Review of Philosophy and Psychology*, published online: 14 June 2018, doi: 10.1007/s13164-018-0400-9.

Feltz A., Cokely E.T. (2008), "The Fragmented Folk: More Evidence of Stable Individual Differences in Moral Judgments and Folk Intuitions," [in:] *Proceedings of the 30th Annual Conference of the Cognitive Science Society*, B.C. Love, K. McRae, V.M. Sloutsky (eds), TX: Cognitive Science Society, Austin: 1771–1776.

Fisher M., Knobe J., Strickland B. et al. (2017), "The Influence of Social Interaction on Intuitions of Objectivity and Subjectivity," *Cognitive Science* 41 (4): 1119–1134.

Forsyth D.R. (1980), "A Taxonomy of Ethical Ideologies," *Journal of Personality and Social Psychology* 39 (1): 175–184.

Gert J. (2013), "Color Constancy and Dispositionalism," *Philosophical Studies* 162 (2): 183–200.

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.

Goodwin G.P., Darley J.M. (2008), "The Psychology of Meta-Ethics: Exploring Objectivism," *Cognition* 106 (3): 1339–1366.

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.

Heiphetz L., Young L.L. (2017), "Can Only One Person Be Right? The Development of Objectivism and Social Preferences Regarding Widely Shared and Controversial Moral Beliefs," *Cognition* 167: 78–90.

Joyce R. (2015a), "Moral Objectivity and Moral Relativism," [in:] *Stanford Encyclopedia of Philosophy*, E.N. Zalta (ed.), URL = https://plato.stanford.edu/entries/moral-anti--realism/moral-objectivity-relativism.html [Accessed 2.10.2019].

Joyce R. (2015b), "Moral Anti-Realism," [in:] *Stanford Encyclopedia of Philosophy*, E.N. Zalta (ed.), URL = https://plato.stanford.edu/entries/moral-anti-realism/ [Accessed 2.10.2019].

Gowans C. (2015), "Moral Relativism," [in:] *Stanford Encyclopedia of Philosophy*, E.N. Zalta (ed.), URL = https://plato.stanford.edu/entries/moral-relativism/ [Accessed 5.10.2019].

McHugh, M. L. (2012), "Interrater Reliability: The Kappa Statistic," *Biochemia Medica*, 22 (3): 276–282.

Moss D. (2017), "Experimental Philosophy, Folk Metaethics and Qualitative Methods," *Teorema* 36 (3): 185–203.

Nichols S. (2004), "After Objectivity: An Empirical Study of Moral Judgment," *Philosophical Psychology* 17 (1): 3–26.

Pölzler T. (2018), "How to Measure Moral Realism," *Review of Philosophy and Psychology* 9 (3): 647–670.

Pölzler T., Wright J.C. (2020), "Anti-Realist Pluralism: A New Approach to Folk Metaethics," *Review of Philosophy and Psychology* 11 (1): 53–82.

Pölzler T., Wright J.C. (2019), "Empirical Research on Folk Moral Objectivism," *Philosophy Compass* 14 (5): e12589.

Rai T.S., Holyoak K.J. (2013), "Exposure to Moral Relativism Compromises Moral Behavior," *Journal of Experimental Social Psychology* 49 (6): 995–1001.

Rose D., Nichols S. (2019), "From Punishment to Universalism," *Mind & Language* 34 (1): 59–72.

Rysiew P. (2011), "Relativism and Contextualism," [in:] *A Companion to Relativism*, S.D. Hales (ed.), Blackwell Publishing, West Sussex (UK): 286–305.

Sarkissian H., Park J., Tien D. et al. (2011), "Folk Moral Relativism," *Mind & Language* 26 (4): 482–505.

Sarkissian H., Phelan M. (2019), "Moral Objectivism and a Punishing God," *Journal of Experimental Social Psychology* 80: 1–7.

Wainryb C., Shaw L.A., Langley M. et al. (2004), "Children's Thinking about Diversity of Belief in the Early School Years: Judgments of Relativism, Tolerance, and Disagreeing Persons," *Child Development* 75 (3): 687–703.

Wright J.C. (2018), "The Fact and Function of Meta-Ethical Pluralism: Examining the Evidence," [in:] *Oxford Studies in Experimental Philosophy* (vol. 2), T. Lombrozo, S. Nichols, J. Knobe (eds), Oxford University Press, Oxford: 119–150.

Wright J.C., Grandjean P.T., McWhite C.B. (2013), "The Meta-Ethical Grounding of Our Moral Beliefs: Evidence for Meta-Ethical Pluralism," *Philosophical Psychology* 26 (3): 336–361.

Yilmaz O., Bahçekapili H.G. (2015), "Without God, Everything is Permitted? The Reciprocal Influence of Religious and Meta-Ethical Beliefs," *Journal of Experimental Social Psychology* 58: 95–100.

Yilmaz O., Bahçekapili H.G. (2018), "Meta-Ethics and the Mortality: Mortality Salience Leads People to Adopt a Less Subjectivist Morality," *Cognition* 179: 171–177.

Young L., Durwin A.J. (2013), "Moral Realism as Moral Motivation: The Impact of Meta-Ethics on Everyday Decision-Making," *Journal of Experimental Social Psychology* 49 (2): 302–306.

Zijlstra L. (2019), "Folk Moral Objectivism and Its Measurement," *Journal of Experimental Social Psychology* 84: 103807.