An empirical argument against moral non-cognitivism

Thomas Pölzler & Jennifer Cole Wright

Published online: 23 Jul 2020.

Submit your article to this journal ↗

View related articles ↗

View Crossmark data ↗

Routledge
Taylor & Francis Group

✇ OPEN ACCESS | Check for updates

An empirical argument against moral non-cognitivism

Thomas Pölzler ◉[a] and Jennifer Cole Wright[b]

[a]Department of Philosophy, University of Graz, Graz, Austria; [b]Department of Psychology, College of Charleston, Charleston, SC, USA

ABSTRACT
According to non-cognitivism, moral sentences and judgements do not aim to represent how things morally are. This paper presents an empirical argument against this view. We begin by showing that non-cognitivism entails the prediction that after some reflection competent ordinary speakers' semantic intuitions favor that moral sentences and judgements do not aim to represent how things morally are. At first sight, this prediction may seem to have been confirmed by previous research on folk metaethics. However, a number of methodological worries lead us to doubt this interpretation. We, therefore, conducted a psychological study that alleviates these worries as far as possible. It turned out that competent ordinary speakers' reflective semantic intuitions dominantly fail to favor that moral sentences do not aim to represent how things morally are. This challenge to non-cognitivism is defended and supplemented by considering deflationary theories of moral truth and middle ground theories in the cognitivism/non-cognitivism debate.

1. Introduction

The practice of morality raises the following two closely related questions in semantics and philosophical psychology: What do moral sentences mean? And what does it mean to make a moral judgement? One influential answer to these questions is provided by moral non-cognitivism (henceforth simply 'non-cognitivism').[1]

CONTACT Thomas Pölzler ✉ thomas.poelzler@uni-graz.at 🖅 Department of Philosophy, University of Graz, Attemsgasse 25/II, 8010 Graz, Austria

[1]As we define non-cognitivism as belonging to moral semantics and philosophical psychology, our considerations are not meant to apply to pragmatic versions of non-cognitivism (according to which

Non-cognitivism is most readily associated with the metaethical theories of Ayer ([1936] 1952), Hare (1952), Blackburn (2000, 2006) and Gibbard (1990, 2003).[2] Its proponents generally agree on the following negative claim: moral sentences and judgements do not aim to represent how things morally are (see Schroeder 2010).[3] Take, for example, the judgement that torturing puppies for fun is wrong. Non-cognitivists claim that by making this judgement we do not aim to represent torturing puppies as having the moral property of wrongness. Rather, morality is supposed to have a more practical purpose. It is closely tied to motivating action. Blackburn, for example, states:

> […] the function of normative sentences is not to represent either peculiar Moorean facts about the world or more mundane empirical facts about the world. It is to avow attitudes, to persuade others, to insist on conformities and prescribe behavior. (Blackburn 2006, 148)

Non-cognitivism is a rather extreme view. If correct, it would have significant implications for our understanding of morality, including its metaphysical and epistemological status (Loeb 2008; Ridge 2013; Schroeder 2010). Nevertheless, it is defended by a sizable minority of philosophers (philpapers.org 2012). This makes providing evidence for or against non-cognitivism important.

How can we assess non-cognitivism? Arguments in moral semantics and philosophical moral psychology have traditionally mainly appealed to *a priori* or non-scientific empirical evidence (see non-cognitivism's main proponents, as cited above, as well as, e.g. Dorr 2002; Geach 1958, 1965). In recent years, however, metaethicists have also increasingly referred to scientific evidence. For example, some non-cognitivists have suggested that their view is supported by the supposition that moral judgements evolved because they performed certain adaptive functions (Gibbard 1990; Kitcher 2005)[4]; and both cognitivists and non-cognitivists have utilized research about moral judgements' current empirical properties as well (such as their relation to emotions or their motivating effects in patients with certain kinds of brain damages) (e.g. Prinz 2006, 2007; Roskies 2003).[5]

---

moral speech acts are not assertions). For a discussion of these different versions of non-cognitivism see Joyce (2007).

[2]Another metaethicist who is often classified as a non-cognitivist is Stevenson. We do not mention him above because his theory fails to be non-cognitivist in a 'pure' sense (see Section 8). Note, furthermore, that some philosophers (e.g. Eriksson 2009) have also denied that Hare was a pure non-cognitivist; and that Blackburn (1996) has declined to be labelled a non-cognitivist himself.

[3]While non-cognitivists deny that moral sentences and judgements aim to represent how things morally are, they do not deny that they can aim to represent how things non-morally are. This will be explained in detail in Section 5.4.

[4]Kitcher later converted to an evolutionary variant of cognitivist naturalism (2011).

In this paper, we attempt to add to this growing body of empirical evidence about non-cognitivism. Our argument draws on a kind of research that has so far been surprisingly neglected by those who accept a scientific approach to the cognitivism/non-cognitivism debate, namely psychological research on folk metaethics (that is, psychological research on what ordinary people think about metaethical issues). We will present novel research of this kind that casts doubt upon non-cognitivism. It suggests that moral sentences actually do aim to represent how things morally are.

Here is how we will proceed. First, we will explain non-cognitivism in more detail (Section 2). On the basis of this characterization, we will then argue that it entails the empirical prediction that after some reflection competent ordinary speakers' semantic intuitions favor that moral sentences and judgements do not aim to represent how things morally are (Section 3). At first sight this prediction may seem to have been confirmed by previous research on folk metaethics. However, a number of methodological worries lead us to doubt this interpretation (Section 4). We conducted a psychological study that alleviates these worries as far as possible. It turned out that competent ordinary speakers' reflective semantic intuitions dominantly do not favor that moral sentences do not aim to represent how things morally are (Section 5). This challenge to non-cognitivism will be defended and supplemented by considering deflationary theories of moral truth (Section 6) and middle ground theories in the cognitivism/non-cognitivism debate (Section 7).

2. Non-cognitivism

In the introduction, we defined non-cognitivism as the view that moral sentences and judgements do not aim to represent how things morally are. This claim has important implications for morality's mental exemplification and truth-aptness.

2.1. Denial of cognitive involvement

Intentional mental states can be classified as either cognitive (belief-like) or non-cognitive (desire-like) (e.g. Platts 1979; Smith 1994).[6] Whether a state falls in one or the other of these categories is typically taken to be

---

[5]In terms of its metaphysical and epistemological implications, non-cognitivism may also be claimed to be supported by more well-known science-based arguments, such as the argument from moral disagreement and evolutionary debunking arguments (even though these arguments do not directly or exclusively attempt to support non-cognitivism).

[6]In line with some research in this area, the terms 'belief' and 'desire' are used in a very broad sense here.

determined by its functional role. Cognitive states are defined as aiming at correctly representing the world; as mapping how things are (*mind-to-world* direction of fit). Non-cognitive states, on the other hand, are rather thought to lead those who have them to change the world (*world-to-mind* direction of fit).[7]

Given this taxonomy of intentional mental states, non-cognitivism's claim that moral judgements and sentences do not aim to represent how things morally are entails a view about these sentences' and judgements' related mental states. To say or think that something is morally right, wrong, good, bad, and so on cannot involve expressing or having a cognitive mental state with content representing (mapping onto) any particular moral facts. Non-cognitivists are rather committed to the view that these sentences and judgements express or are constituted by non-cognitive mental states.

The particular nature of the relevant non-cognitive mental states is contested. For example, non-cognitivists have claimed that moral sentences express feelings of approval or disapproval towards their objects (Ayer [1936] 1952), intentions (Hare 1952)[8], dispositions to feel sentiments of approval or disapproval (Blackburn 2000) or the acceptance of norms that require or permit us to have certain emotions (Gibbard 1990). That said, in whatever way the relevant mental states are characterized, non-cognitivists all agree that these states aim at changing rather than representing how things are.

2.2. Denial of truth-aptness

Moral truth is most commonly understood in a 'correspondence-theoretic' sense (see, e.g. Huemer 2005; Miller 2009; Sayre-McCord 2015). In this sense, a moral sentence or judgement is true if and only if it correctly represents how things are. For example, the sentence 'Torturing puppies for fun is wrong' is true if and only if it is a fact that torturing puppies for fun is wrong.

Non-cognitivists deny that moral sentences and judgements aim to represent how things morally are. But if these sentences and judgements do

---

[7]Elizabeth Anscombe has provided the most famous illustration of the belief-desire model (1963). Suppose a person is shopping for groceries. Beliefs, Anscombe claims, can be compared to a detective that keeps track of this person. The detective aims at representing what the person puts into her cart. Desires, on the other hand, can be compared to the person's grocery list. The list tells her what to put in the cart, based on what she wants to buy.

[8]Hare did not explicitly specify the kind of non-cognitive mental states that he took moral sentences to express. For the above interpretation, see van Roojen (2013).

not aim to represent moral facts at all then, obviously, they cannot represent these facts correctly or incorrectly. Given the correspondence theory of moral truth, non-cognitivists are thus committed to the view that moral sentences and judgements are not truth-apt; that is, that they are neither true nor false.

Denying moral truth-aptness raises a number of theoretical challenges (e.g. Geach 1958, 1965). Contemporary non-cognitivists have therefore typically insisted that while moral sentences and judgements are not truth-apt in a correspondence-theoretic sense, they are truth-apt in an alternative deflationary sense. The basic idea of deflationism is that to say that it is 'true' that something is morally right, wrong, good, bad, and so on just means to say that one affirms this sentence or judgement (Blackburn 2000; Gibbard 2003). For example, by saying that it is true that torturing puppies for fun is wrong one merely affirms one's commitment to the sentence 'Torturing puppies for fun is wrong'.

3. Empirical predictions

Now that we have some understanding of non-cognitivism, let us turn to the question of its empirical predictions; in particular, its predictions with regard to folk metaethics.

The nature of non-cognitivism's predictions depends significantly on whether one takes an internalist or externalist approach to moral semantics and philosophical moral psychology. In this paper we assume an internalist approach (as defended, e.g. by Finlay 2008; Laskowski and Finlay 2017). On this approach, the meaning of moral sentences and judgements is fixed by factors that are internal to ordinary speakers' minds. In particular, it is fixed by these speakers' *intuitions*, by which we mean their pre-theoretical dispositions to apply or refrain from applying concepts. Claims about the meaning of moral sentences and judgements are justified to the extent that they conform to these intuitions (Jackson 1998; Kauppinen 2007; Loeb 2008).

Such an internalist approach to semantics and philosophical psychology may not be plausible with regard to all subject matters. For example, Putnam (1975) famously argued that the fact that our ordinary concept of water refers to H_2O is partly determined by factors outside of people's minds, namely by water's actual nature (chemical structure). However, there are good reasons to believe that internalism is preferable when it comes to morality. Moral properties, if such things exist, lack most or all features that have traditionally been claimed to favor externalist

analyses. For example, relating to Putnam's above case, these properties could not be as easily identified and pointed to as water, and without knowledge of their nature (see Finlay 2008; see also Laskowski and Finlay 2017).[9]

Given an internalist approach, non-cognitivism's empirical predictions might seem trivial. The view simply seems to imply that ordinary speakers must be disposed to apply the terms 'moral sentence' and 'moral judgement' only to a kind of sentences and judgements that do not aim to represent how things morally are. But the situation is actually more complicated. Critics have put forward a number of objections against (internalism-based) scientifically informed arguments in moral semantics and philosophical moral psychology (e.g. Kauppinen 2007; Ludwig 2007, 2010). While these objections fail to undermine such arguments in the way their proponents have claimed (e.g. Nadelhoffer and Nahmias 2007; Sytsma and Livengood 2015), they do suggest that ordinary speakers' intuitions must fulfill three additional conditions in order to be relevant to moral semantics and philosophical moral psychology (see Hannon 2018; Pölzler 2018a, 2018c).

REFLECTION: The question of morality's representative function is abstract and complex. In their everyday lives, ordinary speakers typically do not think about this question. If a speaker reports his or her immediate intuition the likelihood of underlying misinterpretations is therefore high. What matters for moral semantics and philosophical moral psychology are then not these immediate intuitions but rather speakers' *reflective* intuitions, that is, how speakers are disposed to apply moral concepts after having thoroughly thought about the case at issue (Kauppinen 2007; Ludwig 2007; Sosa 2007, 2009).

COMPETENCY: To be relevant for our purposes speakers' intuitions must also reflect minimal competency with regard to moral terms and concepts (Kauppinen 2007; Ludwig 2007, 2010; Sayre-McCord 2008). This competency must be specified in uncontroversial and theoretically neutral ways. With regard to the cognitivism/non-cognitivism debate, for example, one may discount the intuitions of speakers who fail to understand the difference between normative and metaethical statements

---

[9]Even on externalist approaches ordinary speakers' intuitions might be relevant to the philosophical analysis of moral sentences and judgements. We acknowledge that this possibility is controversial. Yet, it shows that while we consider internalism to be much more plausible, and our arguments are explicitly grounded in internalism, the paper may not (fully) stand and fall with it. The empirical data reported in Section 5 may give rise to an externalist challenge to non-cognitivism as well (see, e.g., Jackman 2005; Loeb 2008).

about morality, or the difference between truth-apt and not truth-apt sentences.

SEMANTICS: Finally, ordinary speakers' intuitions about cases may sometimes be explained by their assumptions about the context of these cases or the intentions of characters (e.g. Kauppinen 2007). For example, a speaker might only respond that in a case of moral disagreement both parties are right because s/he assumes that one of the parties intends to provoke the other, and does not really hold the judgement that s/he expresses. Intuitions that are grounded in such pragmatic considerations must be discounted as well. Only *semantic* intuitions count as evidence about the meaning of moral sentences and judgements; that is, only intuitions that essentially are about this meaning.

In sum, then, a closer look suggests that even on the assumption of semantic internalism ordinary speakers need not generally have the intuition that moral sentences and judgements do not aim to represent how things morally are in order for non-cognitivism to be supported. The view only entails that *competent* ordinary speakers have a *semantic* intuition with this content *after some reflection*.

> *Non-Cognitivism's Empirical Prediction with regard to Folk Metaethics*: After some reflection, competent ordinary speakers have the semantic intuition that moral sentences and judgements do not aim to represent how things morally are (that these sentences and judgements do not express or are not constituted by cognitive mental states, and are not truth-apt in a correspondence-theoretic sense).

4. Previous research

In recent years more and more psychologists have begun to investigate folk intuitions about moral objectivity (in the sense of morality's independence from the mental states of observers) (e.g. Beebe 2014, 2015; Beebe and Sackris 2016; Beebe et al. 2015; Fisher et al. 2017; Goodwin and Darley 2008, 2010, 2012; Nichols 2004; Sarkissian et al. 2011; Wright 2018; Wright, Grandjean, and McWhite 2013, 2014). At first sight, some of these studies seem to allow conclusions about non-cognitivism's above prediction. However, we believe that this impression is misleading.

4.1. Methods and results

Most studies on folk moral objectivity have asked participants to interpret cases of moral disagreement. In their extant forms these measures have not been fine-grained enough to cover non-cognitivism as a specific

variant of anti-objectivism, and thus do not provide information about the distribution of cognitivist versus non-cognitivist intuitions (see Pölzler 2018b). Participants in Sarkissian et al.'s study, for example, were presented cases in which two parties disagree with each other about a moral issue and then asked to rate how much they agree with the statement that 'at least one of them [the disagreeing parties] must be wrong' (2011, 487). But rejections of this statement cannot only be explained by participants holding that there is no such thing as being right or wrong with regard to the moral question at issue (non-cognitivism); they can also be explained by their believing that both parties in this disagreement can be right (subjectivism).

In addition, some researchers have tried to determine people's intuitions about moral objectivity in an alternative way as well, namely by means of truth-aptness measures (e.g. Goodwin and Darley 2008, 2012; Wright 2018; Wright, Grandjean, and McWhite 2013, 2014). With the exception of Wright's (2018) study, which involved different questions, these measures have been adopted from Goodwin and Darley (2008):

> How would you regard the previous statement? Circle the number. (1) True statement. (2) False statement. (3) An opinion or attitude. (Goodwin and Darley 2008, 1344)

> According to you, can there be a correct answer as to whether this statement is true?
>   [yes]
>   [no] (Goodwin and Darley 2008, 1351)

Researchers have interpreted (3) and [no] responses as indicating subjectivism. One might argue, however, that these responses rather entail non-cognitivism (Pölzler 2017). If truth-aptness-based studies were reinterpreted in this way, all of them would either provide evidence for non-cognitivism's empirical prediction about folk metaethics or would at worst yield a mixed picture. In Goodwin and Darley's (2008) study, for example, 62% of participants chose the 'opinion or attitude' option (experiment 1), and 47% chose the 'no' option (experiment 2) – which would have to be interpreted as meaning that 62% and 47% of respondents have non-cognitivist intuitions.

But is this reinterpretation appropriate? In our view, there is reason to doubt that even these truth-aptness-based studies on folk moral objectivity provide valid information about the distribution of cognitivist and non-

cognitivist intuitions, at least in the sense that would be required for arguments in moral semantics and philosophical moral psychology.

4.2. Worries

A first worry concerns the wording of the above truth-aptness measures. In our pilot study participants' verbal explanations suggested that if unexplained, many of them understand the term 'truth' in an objectivist sense. But for a moral sentence to be true cannot only mean that it correctly represents an objective moral fact; it can also mean that it correctly represents a subjective moral fact (such as a fact about the moral beliefs of individuals or cultures). Moreover, Goodwin and Darley's first measure may also have exaggerated the proportion of non-cognitivist responses by involving the term 'opinion'. This term cannot only denote a not truth-apt sentence but also truth-apt sentences; in particular, truth-apt sentences that one is somewhat uncertain about (such as in 'In my opinion, Donald Trump will not be reelected') (Beebe 2015, 13–14).

To varying degrees, the above studies are also subject to more general methodological worries that one of us has raised elsewhere (Pölzler 2017, 2018b, 2018c). First, some of the studies did not test whether participants took the sentences that they were presented with to be moral. So even where they sided with non-cognitivism it sometimes may not have been *moral* non-cognitivism. Second, none of the studies made sure that participants understood moral truth in a correspondence-theoretic way. But if they held some alternative (e.g. deflationary) account, then the studies' responses would not straightforwardly indicate intuitions about the cognitivism/non-cognitivism distinction, as introduced above. And third, despite some validity checks, one may also plausibly worry that participants systematically conflated moral truth-aptness with distinct issues (such as perceptions of consensus), or did not have any intuitions about this matter at all.

Finally, the above research on folk moral objectivity also did not fully account for REFLECTION, COMPETENCY and SEMANTICS. The study that comes closest to doing so is Wright's (2018). Before presenting her main tasks she provided participants with detailed explanations of the truth-apt/not truth-apt distinction and tested their understanding of these explanations. However, as one of us has argued elsewhere (Pölzler 2018b, 2018c), Wright's explanations may have introduced a bias in favor of non-cognitivism (by linking non-cognitivism to disagreement,

and providing controversial and one-sided examples of not truth-apt sentences).

To be fair, the extent to which any of the above worries applies is an empirical question. It is possible that participants would have responded in the same way if the term 'opinion' had not been used in the non-cognitivist answer option, if it had been made sure that they do not conflate moral truth-aptness with distinct issues, if they had engaged in sufficient reflection, and so on. But as long as the above plausible alternative explanations have not been ruled out studies on folk moral objectivity nevertheless cannot be claimed to provide significant evidence about non-cognitivism's empirical prediction.

5. New study

In response to worries such as those mentioned in the previous section, we conducted a comprehensive study on folk moral objectivity that was based on an improved experimental design (Pölzler and Wright 2020). In our view, some results of this study provide valid evidence about non-cognitivism's prediction that competent ordinary speakers have the reflective semantic intuition that moral sentences and judgements do not aim to represent how things morally are.

5.1. Participants

Our online survey initially involved 172 participants. To decrease insufficient effort responding (Pölzler, forthcoming), and to conform to REFLECTION, COMPETENCY and SEMANTICS (Section 3), 55 of these participants (32%) were excluded on the basis of the following criteria: (1) failure in one of several attention checks, (2) overly fast completion of the study, (3) bad performance in comprehension checks, (4) confused or irrelevant verbal explanations of their responses.[10] This left us with 117 participants. Of these, 67 were from Amazon Mechanical Turk and 50 from the College of Charleston. 63% were female; their age varied between 18–64 years (M = 29.6); and they were 86% Caucasian, 4% African American, 4% Asian American, 3% Hispanic, and 3% other.

---

[10]Failure in any of our attention checks led to participants being automatically excluded from the study. In contrast, bad performance with regard to the completion time, comprehension check and verbal explanations requirements were only seen as a *prima facie* reason for exclusion. For example, if a participant completed the study very fast but did reasonably well in all or most comprehension checks and verbal explanation we nevertheless kept his/her data set.

5.2. Methods

Our study involved a number of abstract and concrete tasks that were ultimately designed to capture intuitions about moral objectivity. In what follows we focus on those tasks that were targeted at the cognitivism/non-cognitivism debate: our abstract and concrete truth-aptness tasks.

At the beginning of our study, we motivated participants to engage in reflection. For example, we informed them that they 'cannot speed through' the study, that the study involves various comprehension and attention checks, and that we are only looking for people 'who will be serious and conscientious about reading through and answering the questions carefully and honestly'. Then participants received a detailed explanation about the distinction between metaethics and normative ethics and were informed that our study is exclusively about metaethics. Their understanding of this explanation was tested by both a theoretical and a classification exercise.

Before their first truth-aptness task participants also received a detailed explanation about the distinction between truth-apt and not truth-apt sentences:

Truth-apt sentences

Truth-apt sentences express beliefs about facts. These facts can be either objective (independent from mental states) or subjective (dependent on what particular individuals, the majority of individuals in a culture, or others think). In any case, in uttering these sentences we make claims about how things are; what is the case.

Sometimes we do not know whether a truth-apt sentence is true or false. However, even when this happens, the sentence is *still* truth-apt. The sentence is intended to make a claim about how things are – we just don't know whether it does so successfully.

Sometimes a truth-apt sentence turns out to be false. But again, even when this happens, the sentence is *still* truth-apt. The sentence is intended to make a claim about how things are – it just does not do so successfully.

Not truth-apt sentences

Other sentences are not truth-apt. In uttering these sentences we do not make claims about how things are; what is the case. We neither express beliefs about objective nor about subjective facts. Instead, not truth-apt sentences express feelings, intentions, emotions or attitudes. Since they do not make claims about how things are they are neither true nor false.

Participants' understanding of this explanation was tested and improved by two kinds of comprehension checks. First, participants had to complete the following theoretical exercise[11]:

Please select all of the following statements that are correct (you can select more than one):

- A false sentence cannot be truth-apt. [NOT CORRECT]
- Truth-apt sentences can express beliefs about facts that are subjective, that is, facts that are determined by the moral beliefs of individuals, the dominant moral beliefs in cultures, and so on. [CORRECT]
- Truth-apt sentences only express feelings, emotions, intentions or attitudes. [NOT CORRECT]
- Even if we do not know whether a sentence is true or false it can still be truth-apt. [CORRECT]

After that participants were asked to classify the following non-moral sentences as either truth-apt or not truth-apt.[12]

- My pencil is sharp. [TRUTH-APT]
- Don't run in the street. [NOT TRUTH-APT]
- She was very sad about what happened. [TRUTH-APT]
- Yikes! [NOT TRUTH-APT]
- Bummer! [NOT TRUTH-APT]
- Have fun storming the castle. [NOT TRUTH-APT]
- Walking in the street is generally safer than running. [TRUTH-APT]
- Garlic lowers cholesterol. [TRUTH-APT]
- Be happy about what happened! [NOT TRUTH-APT]
- John believes that it was fun storming the castle. [TRUTH-APT]

Participants who failed to correctly answer the theoretical question and to classify all sentences correctly were shown the instructions again, and asked to complete the exercise/s that they had failed one more time.

Following these comprehension checks (which were supposed to test and increase reflection and competency), participants received either our abstract or concrete truth-aptness task (the order of these tasks was randomized). The abstract task looked as follows:

Think about moral sentences (sentences that express that something is morally good or bad, right or wrong, virtuous or vicious, and so on). Are these sentences truth-apt or not truth-apt?

- Yes, moral sentences are 'truth-apt' – that is, they intend to express how things are; what is the case (either with regard to the objective world or with regard to what particular individuals, cultures, etc. think about morality). Thus, these sentences are either true or false.

---

[11]The order of the answer options was randomized.
[12]The order of these sentences was randomized.

- No, moral sentences are not 'truth-apt' – that is, they do not intend to express beliefs about objective or subjective facts, but rather only express feelings, emotions, intentions or attitudes. Thus, these sentences are neither true nor false.

Prior to the concrete truth-aptness task participants were asked to rate 10-item statements (see Table 1) as either dominantly moral or dominantly non-moral. We instructed them to 'focus on the information given by the sentence and […] not introduce additional assumptions or details about what happened or may have happened, or why' (which was supposed to decrease pragmatic influences). Then participants were asked to classify all of the sentences that they had rated as dominantly moral as either truth-apt or not-truth-apt:[13]

Consider the following sentences. Here we are not interested in whether you believe these sentences are actually true. Our focus is rather on their truth-aptness.

Are these sentences truth-apt (i.e. they intend to express what is the case about individuals' moral beliefs, culturally dominant moral beliefs or the objective world; and thus, they are either true or false) or not truth-apt (i.e. they intend only to express feelings, emotions, intentions or attitudes, and thus, they are neither true, nor false)?
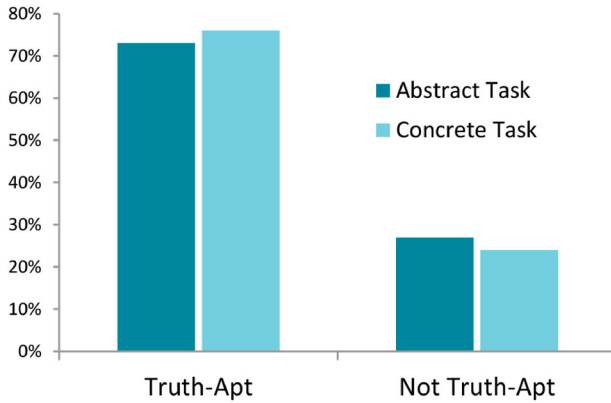
truth-apt

not truth-apt

In addition, participants were also asked to explain their response to the abstract truth-aptness task (immediately after this task), and to explain any inconsistencies between their abstract and concrete responses and within their concrete responses (at a later stage of the study). This was supposed to engage reflection and allow testing the COMPETENCY and SEMANTICS requirements.

5.3. Results

Most participants in our study did reasonably well in the truth-aptness comprehension checks. For example, 81% correctly chose the second and fourth statement in the theoretical exercise at first pass; and 33% correctly classified all sentences in the classification exercise at first pass, with the rest having been mostly able to fully correct their misclassifications at second pass (71%).

---

[13]The order of the sentences was randomized.

Figure 1. Proportion of 'truth-apt' and 'not truth-apt' responses in the abstract and concrete tasks. Abstract task: 73% 'truth-apt' vs. 27% 'not truth-apt'. Concrete tasks: 76% 'truth-apt' vs. 24% 'not truth-apt'.

Both in the abstract and in the concrete tasks the majority of participants' responses favored truth-aptness (Figure 1). In the abstract task 86 participants (73%) stated that, in their view, moral sentences are truth-apt; and 32 participants (27%) stated that they are not truth-apt. This is a greater difference than would be expected by chance (binomial test, $p < .001$). In the concrete task participants' ratings varied with the particular (self-classified moral) statements that they were presented. Overall the item statements prompted 76% 'truth-apt'-responses and 24% 'not truth-apt'-responses. When the frequency of truth-apt vs. not truth-apt classifications was calculated and compared using a paired samples t-test, this showed that participants more frequently classified the item statements as truth-apt than not truth-apt, $t(116) = 9.7$, $p < .001$ (Table 1). These results were also confirmed with a non-parametric Wilcoxon Signed Rank Test, $p < 001$.

Finally, we also considered the intrapersonal consistency of participants' responses, i.e. the degree to which one and the same participant provided identical responses to our tasks. It turned out that the above-reported tendency towards truth-aptness showed some consistency. 41 participants (35%) provided 'truth-apt' responses across 100% of our abstract and concrete tasks. In contrast, only 8 participants (7%) consistently provided 'not truth-apt' responses. Counting participants as consistent if they opted for either 'truth-apt' or 'not truth-apt' in at least 80% of abstract and concrete tasks, the number of consistent participants goes up to 66 (56%). The majority of these participants (57 participants, 86%) again favored truth-aptness.

Table 1. Proportion of 'truth-apt' and 'not truth-apt' responses in the concrete truth-aptness tasks, ordered by size of 'truth-apt' responses.

Moral statement	Truth-apt (%)	Not truth-apt (%)	Binomial test
(1) A world in which wealth is distributed equally is more just than a world in which it is distributed unequally.	86	14	$p < .001$
(2) Before the third month of pregnancy, abortion for any reason is morally impermissible.	85	15	$p < .001$
(3) A country with the death penalty is morally worse than a country without.	84	16	$p < .001$
(4) Consciously discriminating against someone, by not hiring them for a job they are clearly qualified for, just because of their race is not wrong.	82	18	$p < .001$
(5) It is morally reprehensible of the current US administration to limit or prohibit immigration.	81	19	$p < .001$
(6) John cheating (committing adultery) on his wife Elizabeth for no other reason than boredom with his marriage is morally permissible.	78	22	$p < .001$
(7) Men who violently physically punish their children are cruel.	76	24	$p < .001$
(8) Martin Luther King was a righteous man.	66	34	$p = .015$
(9) It was horrendous of Martin Shkreli to overprize drugs that he knew sick people really needed.	64	36	$p = .014$
(10) It is good to do onto others as you would have them do onto you.	62	38	$p = .032$

Notes: Binomial tests indicate that each is a greater difference than would be expected by chance. If we employ the most conservative correction for multiple comparisons (Bonferroni = divide p value by # of comparisons, or .005) – which some believe is too conservative – the top seven statements remain significant.

That said, there were also many participants (58% on our 100% criterion, and 44% on our 80%-criterion) who provided intrapersonally inconsistent responses. They typically regarded most moral statements as truth-apt but some as not truth-apt. How can this inconsistency be explained? Previous studies found that metaethical intuitions often vary with people's strength of agreement to moral sentences (the more strongly they agree with a moral statement, the more objectivist their intuitions about it; e.g. Beebe 2014; Beebe et al. 2015; Goodwin and Darley 2010, 2012) and with people's perceptions of the level of societal consensus (the more widely they believe a moral statement to be accepted by other members of their society, the more objectivist their intuitions about it; e.g. Beebe et al. 2015; Beebe 2014; Goodwin and Darley 2008, 2012). In our study participants' intuitions about truth-aptness were not related with either of these factors ($rs < .1$, $ps > .25$). Given our extensive measures and checks in this regard (see Section 4.1), it is also unlikely that many participants were simply not paying attention or lacking in understanding. Future research hence needs to explore further hypotheses about what explains ordinary speakers' inconsistent intuitions about the truth of moral cognitivism/non-cognitivism.

5.4. Discussion

Non-cognitivism predicts that after some reflection, competent ordinary speakers have the semantic intuition that moral sentences and judgements do not aim to represent how things morally are. Our above study provides evidence about this prediction. Accounting for REFLECTION, COMPETENCY and SEMANTICS, the study's 'not truth-apt' response captures both non-cognitivism's denial of cognitive involvement and its denial of truth-aptness. If non-cognitivism were correct participants should therefore dominantly have been drawn to this response. But our results contradict this prediction. Both in the abstract and in the concrete truth-aptness tasks the broad majority of responses were in favor of truth-aptness (73% in the abstract task and 76% in the concrete tasks). Our study thus presents a challenge to non-cognitivism.

Non-cognitivists may try to meet this challenge in a number of ways. Their two most fundamental potential objections – that participants' responses can be explained by their holding a deflationary theory of moral truth, and that non-cognitivism can be modified so as to account for our results – will be addressed in the following two Sections (Sections 6 and 7). Here we will briefly refute three worries that concern more subtle questions about the interpretation of our results.

First, in our study, we stated the 'not truth-apt' option in terms of moral sentences expressing 'feelings, intentions, emotions or attitudes'. Critics may object that this formulation covers only early non-cognitivism, while leaving out more complex proposals by philosophers such as Blackburn (2000, 2006) and Gibbard (1990, 2003) (who claim that moral sentences express mental states such as sentiments of approval or disapproval, or the non-cognitive acceptance of norms that govern emotions). Suppose we had included these more complex mental states in our 'not truth-apt' option. Might it not have been the case that this option would have been picked far more often?

We do not find this alternative explanation plausible. Getting participants to understand theories such as Blackburn's and Gibbard's may not have been feasible or would at least have required lengthy instructions. Moreover, even in our imperfect formulation participants who tended towards such theories still had plenty of reasons to prefer the 'not truth-apt' over the 'truth-apt' response. First, the mental states that contemporary non-cognitivists appeal to are closely related to or sub-categories of feelings, intentions, emotions or attitudes (see Blackburn 2000; Gibbard 1990). Second, the 'not truth-apt' response also explicitly stated that

moral sentences 'do not intend to express beliefs' (which is accepted by all non-cognitivists, including contemporary ones).[14] And third, our formulation also covers non-cognitivism's implication with regard to truth-aptness, namely that moral sentences are not truth-apt in the sense of correctly or incorrectly representing facts (which is also accepted by all non-cognitivists; for more discussion, see Section 6).

Second, our 'not truth-apt' option states that moral sentences 'do not intend to express beliefs about […] facts'. This description is somewhat inaccurate. Non-cognitivists do not deny that moral sentences state facts; they only deny that they state *moral* facts. For example, even on a non-cognitivist reading our item statement 'John cheating (committing adultery) on his wife Elizabeth for no other reason than boredom with his marriage is morally permissible' states the non-moral fact that John cheated on his wife Elizabeth for no other reason than boredom with his marriage. Thus, non-cognitivists might object that more participants would have gone with the 'not truth-apt' option if we hadn't suggested that it disallows for moral sentences to express beliefs about non-moral facts.

In our view, this explanation is implausible too. Drawing the distinction between stating moral and non-moral facts would have additionally complicated our instructions and answer options. Moreover, it again likely would not have had much of an effect. First, the majority of our item statements (six out of ten) actually did not state non-moral facts at all. For example, the sentence 'Men who violently physically punish their children are cruel' does not state that men in general or any particular men physically punish their children; and the sentence 'Before the third month of pregnancy, abortion for any reason is morally impermissible' does not state that women in general or any particular women have abortions before the third month of pregnancy. Second, participants' responses to statements that do state non-moral facts (statements 5, 6, 8, and 9 in Table 1) and statements that do not state non-moral facts (statements 1, 2, 3, 4, 7 and 10 in Table 1) differed only slightly, with the fact-stating statements being considered only slightly less non-cognitivist (79% cognitivist, 21% non-cognitivist) than the not fact-stating ones (72% cognitivist, 28% non-cognitivist), $t(33) = 2.2$, $p = .032$. As Table 1 shows, all of the individual

---

[14]Some contemporary non-cognitivists hold deflationary theories of beliefs, according to which to say that a moral sentence expresses a belief in a fact is equivalent to just saying this sentence. On this basis, one might claim that even the cognitive involvement part of our options did not give subjects drawn to contemporary non-cognitivist theories a reason to choose "not truth-apt". For a brief discussion of this possibility see Section 6.

statements, including those that do not state non-moral facts, were significantly more cognitivist than would have been expected by chance.

Finally, even though we excluded a considerable proportion of participants prior to analysis, non-cognitivists may object that we still failed to guarantee REFLECTION, COMPETENCY, and SEMANTICS. Participants in our study were excluded on the basis of how they behaved in all tasks of our large study taken together. But a participant who showed sufficient performance and understanding overall might nevertheless have done badly in our truth-aptness tasks in particular. So there is still a (slight) chance that people who engage in more reflection about the truth-apt/not truth-apt distinction, are more competent with regard to it, and are more driven by semantic considerations would dominantly choose the 'not truth-apt'-responses; that is, in other words, there is still a (slight) chance that non-cognitivism manifests itself under more ideal circumstances.

To address this objection we introduced additional exclusion criteria that specifically pertained to participants' performance in and understanding of our truth-aptness tasks. First, we excluded any participant that did not get both the theoretical truth-aptness exercise and all truth-aptness classification exercises right at least at second pass. Second, we independently analyzed participants' verbal explanations of their responses to the abstract truth-aptness task, and excluded any participant whose explanation was classified as confused or irrelevant by both of us. This led us to exclude additional 55 participants (47%), and left us with a remaining sample of 62 'ideal' participants.

Contrary to the above objection, our stricter analysis did not yield a higher proportion of 'not truth-apt' responses. If anything rather the opposite was the case. While not reaching the level of statistical significance, in both the abstract task and the concrete task our ideal participants opted slightly more in favor of truth-aptness, compared to the participants that we had excluded on the basis of our originalcriteria (79% versus 65% in the abstract task, and 78% versus 73% in the concrete task). Moreover, while not analyzable statistically, the ideal sample was also slightly more inclined towards truth-aptness compared to the whole previous sample of 117 participants (79% versus 73% in the abstract task, and 78% versus 76% in the concrete task). This falsifies the hypothesis that increased reflection, competency and semantic considerations lead people in the direction of non-cognitivism. Even under more ideal circumstances, the majority of people seem to side with cognitivism.

While none of the above objections seem particularly powerful there are two worries about our empirical argument against non-cognitivism that deserve more serious consideration. These objections concern participants' understanding of moral truth and potential modifications of non-cognitivism.

6. Moral truth

Suppose a person has the intuition that to say of a moral sentence that it is true simply means to affirm it. Moral sentences can be meaningfully affirmed. For example, it makes sense to affirm that torturing puppies for fun is wrong or that helping others is good. Any person who was committed to such a deflationary theory of moral truth (e.g. Blackburn 2000; Gibbard 2003) should hence be strongly inclined to consider moral sentences truth-apt. Non-cognitivists may argue that this is precisely what happened in our study. Participants did not dominantly choose the 'truth-apt' response because they favored cognitivism; they did so because, being non-cognitivists, they understood moral truth in a deflationary sense.

Is this alternative explanation plausible? We don't think so. Our study's experimental design and results suggest that most 'truth-apt' responses were not accompanied by a deflationist understanding of moral truth; and that even if participants dominantly did have such an understanding this might not have had a large effect on our results.[15]

To begin with, recall that participants in our study were asked to explain their responses to the abstract truth-aptness task (Section 5.2). Only 1 of the 86 participants (1%) who had opted for truth-aptness provided an

---

[15]We also investigated the prevalence of deflationist intuitions as part of our above study and in a follow-up study. However, both of these investigations turned out to be unsuccessful. The first one (in which we asked people to choose between intelligible descriptions of the correspondence and the deflationary theory) sparked lots of confusion, as evinced by participants' verbal explanations. The second study yielded a result that did not allow for any helpful inferences either. This study involved two conditions, SUBSTANTIVE and TRUTH, which each asked for views about the ten moral item statements listed in Table 1. In SUBSTANTIVE participants received tasks of the following kind: "Which of the following statements best fits your view? (1) X is morally m. (2) X is not morally m. (3) X is neither morally y nor not morally m." Tasks in TRUTH had the following form: "Which of the following statements best fits your view? (1) The sentence 'X is morally m' is true. (2) The sentence 'X is morally m' is false. (3) The sentence 'X is morally m' is neither true nor false." Deflationary theory regards the corresponding options in SUBSTANTIVE and TRUTH as equivalent. Any difference between these conditions would hence challenge the hypothesis that participants are drawn towards this theory. However, in our study, we did not find any significant difference between participants' responses in SUBSTANTIVE and TRUTH. This result is compatible with both deflationary theory and correspondence theory. For correspondence theorists will typically also opt for 'It is true that x is morally m' whenever they opt for 'x is morally m' (unless they are fictionalist error theorists or non-cognitivists — and that they are mostly non-cognitivists is precisely what we are challenging).

explanation that suggests adherence to the deflationary theory of moral truth or a protest against our having primed participants with the correspondence theory. In contrast, 54 of these explanations (63%) were formulated in ways that involve *prima facie* correspondence theoretic commitments; in particular, they state (in exactly or very closely related terms) that moral sentences 'express how things are', 'express beliefs', 'have a representational function' or 'state facts' (Table 2).

Here are some examples of explanations that involve *prima facie* commitments to a correspondence-theoretic understanding of moral truth:

- 'I believe that when most people are using moral sentences they are trying to tell you how things are.'
- 'Moral sentences are truth-apt because they express what the particular culture believes to be true.'
- 'I feel like these [moral sentences] would be very matter of fact sentences.'
- 'Moral sentences are truth-apt because you are stating what you think is a fact.'

Moreover, recall that our truth-aptness tasks' answer options were not only stated and explained in terms of non-cognitivism's implications for (correspondence-theoretic) moral truth. We also explained that truth-apt sentences 'express beliefs about facts', and stated our 'not truth-apt' option as saying that moral sentences 'do not intend to express beliefs' (Section 5.2). Non-cognitivists deny that moral sentences express beliefs about moral facts. Hence, even if some non-cognitivist participants in our study assumed the deflationary theory of moral truth they still might not have been (consistently) drawn towards our 'truth-apt' option.

Table 2. Proportions of 'truth-apt' responses that were explained in ways that involve a *prima facie* commitment to the correspondence-theory of moral truth.

Explanation	Proportion (%)
Moral sentences express how things are	13
Moral sentences express beliefs	16
Moral sentences have a representational function	2
Moral sentences state facts	17
Moral sentences express how things are and express beliefs	6
Moral sentences express how things are and state facts	2
Moral sentences express beliefs and state facts	6
Total	63

Note: Participants' explanations do not match these phrases *verbatim* but involve their key terms ('how things are', 'beliefs', 'representation', 'facts').

This option's endorsement of cognitive involvement, and the 'not truth-apt' option's denial of such involvement, may well have led them to go with the latter and to be rightly classified as non-cognitivists.

To both of the above arguments non-cognitivists may respond by making deflationism 'creep' (Dreier 2004) into intuitions about moral 'representation', 'beliefs', 'facts', etc. as well. Isn't it possible that by applying these terms to moral sentences people just mean to affirm them too; for example, that by saying that the sentence 'Torturing puppies for fun is wrong' 'represents reality', 'expresses a belief', or 'states a fact' they ultimately just mean that torturing puppies for fun is wrong (see Wright 1992; Timmons 1998)? If this were the case then the fact that our participants used terms such as 'representation', 'belief', 'fact', etc. in explaining their responses could not be interpreted as a commitment to the correspondence theory of moral truth after all; and the fact that our study's 'truth-apt' option endorsed and our 'not truth-apt' option denied that moral sentences represent reality, express beliefs, state facts, etc. could not have drawn non-cognitivist participants towards the former option either.

However, the hypothesis that many people understand notions such as moral representation, belief, fact, etc. in a deflationary sense is problematic. First, theoretical considerations suggest that deflationism about these terms might diverge very far from ordinary discourse. With regard to non-moral domains, for example, people often acknowledge that an entity represents a certain fact even though this entity does not or cannot affirm the fact (think of thermometers, bat sonar, and human proprioception); and they often affirm propositions without accepting that these propositions represent reality (such as with the proposition that Sherlock Holmes took cocaine). Representation hence does not seem to be treated as equivalent with mere affirmation (see Sinclair 2007).

In our view, creeping deflationism's *prima facie* implausibility shifts the burden of proof. It is not us who need to provide evidence against people assuming deflationary understandings of moral 'representation', 'beliefs', 'facts', etc.; rather, non-cognitivists must provide evidence for this hypothesis.

Second, given the creeping deflationism response, how is one supposed to support the hypothesis that ordinary people understand moral truth in a (non-)deflationary sense at all? Proponents of this response argue that almost any term that people can use to explain their views about moral truth can be read in a deflationary sense. Only specific theoretical commitments, such as about the nature of moral explanations, are

claimed to allow distinguishing between deflationism-based non-cognitivism and correspondence theory-based views. But these theoretical commitments have been spelt out very differently (see, e.g. Dreier 2004 vs. Chrisman 2008), and there is a real danger that creeping deflationism in the end simply leads to a full collapse of the distinction (Tiefensee 2016). Accepting the possibility of deflationism about moral representation, beliefs, facts, etc., hypotheses about whether people tend towards deflationism about moral truth may become largely or completely unfalsifiable – which is a most unwelcome consequence (e.g. Popper [1959] 2002).

Finally, suppose the above arguments failed to extend our challenge to modern deflationism-based versions of non-cognitivism. We would like to stress that even then our results would be significant. They would still present a novel challenge to classic non-cognitivism which is distinct from traditional philosophical objections (such as the Frege-Geach problem; Geach 1958, 1965) as well as from the small number of empirical worries that have so far been raised (such as worries based on research about folk motivational internalism; Strandberg and Björklund 2013).

7. Middle ground theories

A last way in which non-cognitivists might try to meet our challenge is by modifying their view in a way that makes it compatible with our results. In our study, participants dominantly regarded moral sentences as truth-apt. Nevertheless, a noteworthy minority (27%) denied truth-aptness on the abstract level, and half of our participants (50%) denied it with regard to at least one concrete item statement. So cannot non-cognitivists simply move towards some middle ground between cognitivism and non-cognitivism – a middle ground that still preserves most or much of their original position's spirit?

Middle ground theories in the cognitivism/non-cognitivism debate have mainly fallen into one of the following six (partly non-exclusive) categories:

(1) *Variantism*: Some moral sentences and judgements aim to represent how things morally are. Other moral sentences and judgements do not aim to represent how things morally are. That is, cognitivism and non-cognitivism are each correct for non-overlapping parts of moral discourse and thought (Gill 2009; Wright 2018).

(2) *Indeterminism*: Moral sentences and judgements are indeterminate with regard to the cognitivism/non-cognitivism distinction. That is, they are correctly analyzed by both theories (like men with a certain amount of hair may be correctly described both as being bald and as not being bald) (Gill 2009; Joyce 2012).

(3) *Hermeneutic Fictionalism*: Moral sentences aim to represent how things morally are. Moral judgements, in contrast, do not aim to represent how things morally are. That is, cognitivism is correct on the level of moral discourse, and non-cognitivism is correct on the level of moral thought (Kalderon 2005).

(4) *Non-Descriptivist Cognitivism*: Moral sentences do not aim to represent how things morally are. Moral judgements, in contrast, do aim to represent how things morally are. That is, non-cognitivism is correct on the level of moral discourse, and cognitivism is correct on the level of moral thought (Horgan and Timmons 2000; Timmons 1998).

(5) *Hybrid Expressivism*: All moral sentences determinately express both cognitive and non-cognitive mental states (for example, similar to how pejorative slurs express both cognitive and non-cognitive mental states) (e.g. Boisvert 2008; Copp 2001; Ridge 2006).

(6) *Incoherentism*: Moral sentences and judgements determinately are both aiming to represent how things morally are and are not aiming to do so. This renders them incoherent (similar to the notion of a 'round square') (Loeb 2008).

In order for non-cognitivists' above defense to work two conditions must be met: (1) at least one of these middle ground theories must be consistent with our study's results, and (2) this theory must preserve most or much of non-cognitivism's original spirit. In what follows we will argue that at least in their above generic forms no theory actually meets both of these conditions.

To begin with, most of the above theories fit our results rather badly. Take variantism. At first sight this middle ground theory might seem to explain our results pretty well. After all, as pointed out in Section 5.3, many of our participants gave intrapersonally inconsistent responses. They seem to have had the intuition that while some moral sentences aim to represent how things morally are, others do not do so. Note, however, that if variantism were true then at least some of our moral item statements should have prompted *dominantly* 'not truth-apt' responses. Only under these circumstances could an internalist approach provide reasons to interpret these sentences in a non-cognitivist way. But

no statement was considered to be not truth-apt by more than 38% of participants. In fact, most values were significantly lower (see Table 1).[16]

Many versions of indeterminism, hybrid expressivism and incoherentism predict that 'truth-apt' and 'not truth-apt' classifications will be distributed fairly equally for each statement (as each of these sentences is claimed to be indeterminate in terms of their meaning or as involving both cognitivist and non-cognitivist elements). But as just pointed out, in our study, all statements were dominantly classified as 'truth-apt'. The distribution of responses was quite unequal (see Table 1). Non-descriptivist cognitivism, finally, requires that moral sentences dominantly prompt non-cognitivist intuitions. This prediction was falsified most clearly. Our overall proportion of 'truth-apt' responses far outweighed the proportion of 'not truth-apt responses': 73% vs. 27% in the abstract task and 76% vs. 24% in the concrete tasks.

The only middle ground theory that fits our results reasonably well is hermeneutic fictionalism. According to hermeutic fictionalism, moral discourse is to be understood in a cognitivist way, i.e. at aiming to represent how things morally are. This view hence predicts that in a study such as ours moral sentences would dominantly be considered as 'truth-apt' rather than 'not truth-apt' – which is, of course, exactly what we found.

Do the above middle ground theories at least preserve most or much of what motivates non-cognitivist analyses in the first place? Again, for most of these theories, our answer is negative. Non-cognitivism's main attraction has always been that it sidesteps a number of intricate and controversial metaethical questions (Schroeder 2010). If its proponents are right that moral sentences and judgements do not aim to represent moral facts then the question of whether these facts exist does not even arise. No kind of facts in the world can possibly deserve to be called moral (Loeb 2008). As a consequence, non-cognitivists also do not need to explain what moral facts look like, how we can come to know about them, how we can talk about them, and how we can think about them.

But the only middle ground theory that by itself spares its proponents from addressing these metaethical questions is incoherentism. All other theories claim that at least some parts or aspects of moral discourse or thought do aim to represent how things morally are, and in such a

---

[16]Variantists may object that we simply failed to include some of those moral sentences that do not represent how things morally are. This response does not strike us as particularly plausible. After all, we made sure that our item statements are as diverse as possible. They include statements about actions as well as about persons and states of affairs, statements that involve thin as well as statements that involve thick moral concepts, statements that are highly intuitive and statements that are highly counterintuitive, statements that are non-negated and statements that are negated, and so on.

(coherent) way that allows for (some) moral facts to exist. This also holds for hermeneutic fictionalism. Proponents of this theory grant that at least when we *say* that something is morally right, wrong, good, bad, etc. we are purporting to represent a moral fact. In the absence of additional arguments, it could therefore be the case that such facts exist (see Friend 2008), and that hermeneutic fictionalists have to explain what these facts look like, how we can come to know about them, how we can talk about them, and how we can think about them.[17]

Thus, non-cognitivists most likely also cannot meet our challenge by modifying their view in the sense of adopting a middle ground theory.

8. Conclusion

In assessing non-cognitivism various kinds of evidence might be deemed relevant. This paper focused on one particular kind of empirical evidence. First, we argued that non-cognitivism entails the prediction that after some reflection competent ordinary speakers' semantic intuitions favor that moral sentences and judgements do not aim to represent how things morally are. Then we reported a psychological study on folk metaethics that contradicts this prediction, and defended our interpretation against a number of objections. Other kinds of evidence (empirical or *a priori*) may outweigh the proposed reason to reject non-cognitivism. As long as we are only considering ordinary people's intuitions, however, views according to whom all moral sentences determinately have a representative function are at an advantage. Our argument suggests that metaethicists should adopt classic cognitivism (which endorses such an analysis for both moral discourse and moral thought) or hermeneutic fictionalism (which endorses such an analysis for only moral discourse).

---

[17]To reemphasize, the above argument only states that unlike non-cognitivism, hermeneutic fictionalism (in its above generic form) does not by itself entail the non-existence of moral facts; that it is *compatible* with there being such facts. We do not mean to suggest that hermeneutic fictionalism by itself involves a commitment to moral facts. In fact, Kalderon (2005) himself explicitly argues against such a commitment.

Disclosure statement

Funding

ORCID

Thomas Pölzler 🔟 http://orcid.org/0000-0002-4311-0897

References

Anscombe, Gertrude E. 1963. *Intention*. Oxford: Basil Blackwell.

Ayer, Alfred J. [1936] 1952. *Language, Truth and Logic*. New York: Dover.

Beebe, James R. 2014. "How Different Kinds of Disagreement Impact Folk Metaethical Judgments." In *Advances in Experimental Moral Psychology: Affect, Character, and Commitments*, edited by Jennifer C Wright, and Hagop Sarkissian, 167–178. London: Bloomsbury.

Beebe, James R. 2015. "The Empirical Study of Folk Metaethics." *Etyka* 15: 11–28.

Beebe, James R., Runya Qiaoan, Tomasz Wysocki, and Miguel A Endara. 2015. "Moral Objectivism in Cross-Cultural Perspective." *Journal of Cognition and Culture* 15: 386–401.

Beebe, James R., and David Sackris. 2016. "Moral Objectivism Across the Lifespan." *Philosophical Psychology* 29 (6): 919–929.

Blackburn, Simon. 1996. "Commentary on Ronald Dworkins' 'Objectivity and Truth: You'd Better Believe it'." Brown Electronic Article Review Service in Moral and Political Philosophy.

Blackburn, Simon. 2000. *Ruling Passions: A Theory of Practical Reasoning*. Oxford: Oxford University Press.

Blackburn, Simon. 2006. "Antirealist Expressivism and Quasi-Realism." In *The Oxford Handbook of Ethical Theory*, edited by David Copp, 146–162. Oxford: Oxford University Press.

Boisvert, Daniel R. 2008. "Expressive-Assertivism." *Pacific Philosophical Quarterly* 89 (2): 169–203.

Chrisman, Matthew. 2008. "Expressivism, Inferentialism, and Saving the Debate." *Philosophy and Phenomenological Research* 77 (2): 334–358.

Copp, David. 2001. ": Realist-Expressivism: A Neglected Option for Moral Realism." *Social Philosophy and Policy* 18 (2): 1–43.

Dorr, Cian. 2002. "Noncognitivism and Wishful Thinking." *Nous (Detroit, Mich)* 36 (1): 97–103.

Dreier, James. 2004. "Meta-Ethics and the Problem of Creeping Minimalism." *Philosophical Perspectives* 18: 23–44.

Eriksson, John. 2009. ": Homage to Hare: Ecumanism and the Frege-Geach Problem." *Ethics* 120 (1): 8–35.

Finlay, Stephen. 2008. "The Error in the Error Theory." *Australasian Journal of Philosophy* 86 (3): 347–369.

Fisher, Matthew, Joshua Knobe, Brent Strickland, and Frank C Keil. 2017. "The Influence of Social Interaction on Intuitions of Objectivity and Subjectivity." *Cognitive Science* 41 (4): 1119–1134.

Friend, Stacy. 2008. "Hermeneutic Moral Fictionalism as an Anti-Realist Strategy." *Philosophical Books* 49 (1): 14–22.

Geach, Peter T. 1958. "Imperative and Deontic Logic." *Analysis* 18 (3): 49–56.

Geach, Peter T. 1965. "Assertion." *Philosophical Review* 74 (4): 449–465.

Gibbard, Allan. 1990. *Wise Choices, Apt Feelings: A Theory of Normative Judgement*. Cambridge, MA: Harvard University Press.

Gibbard, Allan. 2003. *Thinking How to Live*. Cambridge, MA: Harvard University Press.

Gill, Michael B. 2009. "Indeterminacy and Variability in Meta-Ethics." *Philosophical Studies* 145 (2): 215–234.

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

Goodwin, Geoffrey P., and John M Darley. 2010. ": The Perceived Objectivity of Ethical Beliefs: Psychological Findings and Implications for Public Policy." *Review of Philosophy and Psychology* 1 (2): 161–188.

Goodwin, Geoffrey P., and John M Darley. 2012. "Why Are Some Moral Beliefs Perceived to be More Objective Than Others?" *Journal of Experimental Social Psychology* 48 (2012): 250–256.

Hannon, Michael. 2018. "Intuitions, Reflective Judgments, and Experimental Philosophy." *Synthese* 195 (19): 4147–4168.

Hare, Richard M. 1952. *The Language of Morals*. Oxford: Oxford University Press.

Horgan, Terry, and Mark Timmons. 2000. "Nondescriptivist Cognitivism: Framework for a New Metaethic." *Philosophical Papers* 29 (2): 121–153.

Huemer, Michael. 2005. *Ethical Intuitionism*. Basingstoke: Palgrave MacMillan.

Jackman, Henry. 2005. "Intuitions and Semantic Theory." *Metaphilosophy* 36 (3): 363–380.

Jackson, Frank. 1998. *From Metaphysics to Ethics: A Defense of Conceptual Analysis*. Oxford: Clarendon Press.

Joyce, Richard. 2007. "Moral Anti-Realism." In *The Stanford Encyclopedia of Philosophy*, edited by Edward N Zalta. http://plato.stanford.edu/entries/moral-anti-realism/

Joyce, Richard. 2012. "Metaethical Pluralism: How Both Moral Naturalism and Moral Skepticism may be Permissible Positions." In *Ethical Naturalism: Current Debates*, edited by Susana Nuccetelli, and Gary Seay, 89–109. Cambridge: Cambridge University Press.

Kalderon, Mark E. 2005. *Moral Fictionalism*. Oxford: Oxford University Press.

Kauppinen, Antti. 2007. "The Rise and Fall of Experimental Philosophy." *Philosophical Explorations* 10 (2): 95–118.

Kitcher, Philip. 2005. "Biology and Ethics." In *The Oxford Handbook of Ethical Theory*, edited by David Copp, 163–185. Oxford: Oxford University Press.

Kitcher, Philip. 2011. *The Ethical Project*. Cambridge, MA: Harvard University Press.

Laskowski, Nicholas, and Stephen Finlay. 2017. "Conceptual Analysis in Metaethics." In *Routledge Handbook of Metaethics*, edited by Tristram McPherson and David Plunkett, 536–551. New York: Routledge.

Loeb, Don. 2008. "Moral Incoherentism: How to Pull a Metaphysical Rabbit out of a Semantic Hat." In *Moral Psychology, Vol.2: The Cognitive Science of Morality: Intuition and Diversity*, edited by Walter Sinnott-Armstrong, 355–386. Cambridge: MIT Press.

Ludwig, Kirk. 2007. ": The Epistemology of Thought Experiments: First Person Versus Third Person Approaches." *Midwest Studies in Philosophy* 31: 128–159.

Ludwig, Kirk. 2010. "Intuitions and Relativity." *Philosophical Psychology* 23 (4): 427–445.

Miller, Christian B. 2009. "The Conditions of Moral Realism." *Journal of Philosophical Research* 34: 123–155.

Nadelhoffer, Thomas, and Eddy Nahmias. 2007. "The Past and Future of Experimental Philosophy." *Philosophical Explorations* 10 (2): 123–149.

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

philpapers.org. 2012. *The PhilPapers Surveys: Preliminary Survey Results*. http://philpapers.org/surveys/results.pl.

Platts, Mark. 1979. *Ways of Meaning*. London: Routledge & Kegan Paul.

Pölzler, Thomas. 2017. "Revisiting Folk Moral Realism." *Review of Philosophy & Psychology* 8 (2): 455–476.

Pölzler, Thomas. 2018a. "Can the Empirical Sciences Contribute to the Moral Realism/Anti-Realism Debate?" *Synthese* 195 (11): 4907–4930.

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

Pölzler, Thomas. 2018c. *Moral Reality and the Empirical Sciences*. New York: Routledge.

Pölzler, Thomas. Forthcoming. "Insufficient Effort Responding in Experimental Philosophy." In *Oxford Studies in Experimental Philosophy, Volume 4*, edited by Tania Lombrozo, Joshua Knobe, and Shaun Nichols. Oxford: Oxford University Press.

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

Popper, Karl. (1959) 2002. *The Logic of Scientific Discovery*. New York: Routledge.

Prinz, Jesse J. 2006. "The Emotional Basis of Moral Judgements." *Philosophical Explorations* 9 (1): 29–43.

Prinz, Jesse J. 2007. *The Emotional Construction of Morals*. Oxford: Oxford University Press.

Putnam, Hilary. 1975. "The Meaning of Meaning." In *Mind, Language and Reality: Philosophical Papers, Volume 2*, edited by Hilary Putnam, 215–271. Cambridge: Cambridge University Press.

Ridge, Michael. 2006. ": Ecumenical Expressivism: Finessing Frege." *Ethics* 116 (2): 302–336.

Ridge, Michael. 2013. "Non-Cognitivism." In *International Ency-clopedia of Ethics*, edited by Hugh LaFollette, 3610–3625. Oxford: Wiley-Blackwell.

Roskies, Adina. 2003. "Are Ethical Judgments Intrinsically Motivational? Lessons from "Acquired Sociopathy"." *Philosophical Psychology* 16 (1): 51–66.

Sarkissian, Hagop, David Tien, Jennifer Wright, and Joshua Knobe. 2011. "Folk Moral Relativism." *Mind and Language* 26 (4): 482–505.

Sayre-McCord, Geoffrey. 2008. "Moral Semantics and Empirical Enquiry." In *Moral Psychology, Vol. 2: The Cognitive Science of Morality: Intuition and Diversity*, edited by Walter Sinnott-Armstrong, 403–412. Cambridge: MIT Press.

Sayre-McCord, Geoffrey. 2015. "Moral Realism." In *The Stanford Encyclopedia of Philosophy*, edited by Edward N Zalta. http://plato.stanford.edu/entries/moral-realism/

Schroeder, Mark. 2010. *Noncognitivism in Ethics*. New York: Routledge.

Sinclair, Neil. 2007. "Propositional Clothing and Belief." *The Philosophical Quarterly* 57 (228): 342–362.

Smith, Michael. 1994. *The Moral Problem*. Oxford: Blackwell.

Sosa, Ernest. 2007. "Experimental Philosophy and Philosophical Intuition." *Philosophical Studies* 132 (1): 99–107.

Sosa, Ernest. 2009. "A Defense of the Use of Intuitions in Philosophy." In *Stich and His Critics*, edited by Michael J Bishop, and Dominic Murphy, 101–112. Oxford: Blackwell Publishing.

Strandberg, Caj, and Fredrik Björklund. 2013. "Is Moral Internalism Supported by Folk Intuitions?" *Philosophical Psychology* 26 (3): 319–335.

Sytsma, Justin, and Jonathan Livengood. 2015. *The Theory and Practice of Experimental Philosophy*. Peterborough: Broadview.

Tiefensee, Christine. 2016. "Inferentialist Metaethics, Bifurcations and Ontological Commitment." *Philosophical Studies* 173 (9): 2437–2459.

Timmons, Mark. 1998. *Morality Without Foundations*. Oxford: Oxford University Press.

van Roojen, Mark. 2013. "Moral Cognitivism vs. Non-Cognitivism." In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta. http://plato.stanford.edu/entries/moral-cognitivism/.

Wright, Crispin J. G. 1992. *Truth and Objectivity*. Cambridge: Harvard University Press.

Wright, Jennifer C. 2018. "The Fact and Function of Meta-Ethical Pluralism: Exploring the Evidence." In *Oxford Studies in Experimental Philosophy, Vol. 2*, edited by Tania Lombrozo, Joshua Knobe, and Shaun Nichols, 119–150. Oxford: Oxford University Press.

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

Wright, Jennifer C., Cullen B. McWhite, and Piper T Grandjean. 2014. "The Cognitive Mechanisms of Intolerance: Do Our Meta-Ethical Commitments Matter?" In *Oxford Studies in Experimental Philosophy, Vol. 1*, edited by Joshua Knobe, Tania Lombrozo, and Shaun Nichols, 28–61. Oxford: Oxford University Press.