

Moral Cognitivism *Hallvard Lillehammer*

Abstract: The paper explicates a set of criteria the joint satisfaction of which is taken to qualify moral judgements as cognitive. The paper examines evidence that some moral judgements meet these criteria, and relates the resulting conception of moral judgements to ongoing controversies about cognitivism in ethics.

I. Introduction

There is much disagreement about the cognitive status of moral judgements. The main interest of this disagreement does not derive from the aim of producing a theory of mental states, but rather from concerns about the consequences of the answer for the status of moral claims. The problem arises because it is unclear how moral judgements can be better or worse, true or false, if they are to be identified with non-cognitive states like emotions, desires or preferences. Yet from a common-sense perspective, moral judgements are arguably better or worse, true or false. We would like a theory of moral judgements to explain how this can be so. Despite valiant attempts by non-cognitivists to save the appearances, the prevailing view remains that only cognitivism can supply the goods.¹ Nevertheless, there is no agreement among cognitivists and their opponents about what it takes for judgements to be cognitive.² Yet without a clear idea of what makes judgements cognitive we do not understand what the debate over cognitivism in ethics is about.

In this paper I explicate one conception of what it takes for moral judgements to be cognitive. I then examine the hypothesis that moral judgements are cognitive so understood. By the term 'judgement' I refer

1 Cf. Brink (1989), Blackburn (1998, 1998), Smith (1994), Horgan and Timmons (2000).

2 Cf. Wiggins (1990), Horgan and Timmons (2000).

Address for correspondence: *Philosophical Papers*, Department of Philosophy, Rhodes University Box 94, Grahamstown 6140, South Africa

Philosophical.Papers@ru.ac.za (Managing Editor)

W.Jones@ru.ac.za (Editor)

<http://www.rhodes.ac.za/philosophy/PhilosophicalPapers/>

Cover and text design: Graunt Kruger

Website design: Francis Williamson

International Board of Consultants

Lynne Rudder Baker

(Massachusetts)

Jonathan Barnes (Genève)

Lars Bergström (Stockholm)

Tyler Burge (California—Los Angeles)

David Chalmers (Arizona)

Andy Clark (Sussex)

Lorraine Code (York, Canada)

Emmanuel Chukwudi Eze (DePaul)

Terrence Horgan (Memphis)

Jennifer Hornsby (Birkbeck College London)

Frank Jackson (Australian National University)

Jaegwon Kim (Brown)

Mark Leon (Witwatersrand)

Michael Levine (Western Australia)

John McDowell (Pittsburgh)

Brian McLaughlin (Rutgers)

Chris Megone (Leeds)

Seumas Miller (Charles Stuart)

Ruth Garrett Millikan (Connecticut)

Michael Pendlebury (Witwatersrand)

John Perry (Stanford)

Philip Pettit (Australian National University)

Amélie Oksenberg Rorty (Brandeis)

Mark Sainsbury (Texas—Austin)

John Searle (California Berkeley)

Elliot Sober (Wisconsin—Madison)

Ernest Sosa (Brown)

Kim Sterelny (Victoria—Wellington & ANU)

Torbjörn Tännsjö (Göteborg)

Susan Wolf (North Carolina)

Crispin Wright (St. Andrews)

Eddy Zemach (Hebrew University of Jerusalem)

to a set of mental states, the contents of which can be expressed by a declarative sentence. If Pepe judges that his inbox is not his postbox, his judgement is a mental state with content expressible by the indicative sentence 'My inbox is not my postbox'. By the term 'cognitive' I refer to a subset of mental states exhibiting marks I shall take to be individually necessary and jointly sufficient for the possession of what I call *representational purport*. Mental states have representational purport if and only if they possess two distinct features, both of which I characterise further below. First, their standard linguistic articulation gives rise to *assertoric contents* the inferential relations between which can be represented by a truth conditional semantics. Second, the mental states themselves are *non-permissively world-guided*, in the sense that their function is partly characterised as aiming to correctly reflect their putative domain. In addition to these jointly sufficient conditions for cognitive status, I consider a third mark of cognitive status appearing in arguments about cognitivism in ethics, namely *motivational detachment*. A mental state is motivationally detached just in case there is no interesting conceptual connection between being in that state and being motivated a certain way. I claim that while motivational detachment is a mark of some moral judgements, it is neither necessary nor sufficient for cognitive status. In closing, I address a frequently raised problem for moral cognitivism, namely that the motivational detachment exhibited by some cognitive states is inconsistent with the practicality of moral judgements. I claim that this problem is illusory.

Cognitivism as here understood is the view that moral judgements have representational purport. I shall not argue that only judgements with representational purport are cognitive. I shall take representational purport to be a sufficient, not a necessary condition for cognitive status. The aim of this strategy is to motivate the view that some moral judgements are cognitive, not to argue that this is the only way for judgements to be cognitive. The hypothesis is that the marks of

representational purport pick out a set of cognitive judgements to which moral judgements belong.³

As defined above, the term 'judgement' picks out a bigger set of mental states than is normally associated with the word for philosophical purposes. Having content expressible in terms of a declarative sentence is a property shared by all states commonly referred to as the propositional attitudes. It is common to think that if Pepe desires to become a computer hobbyist, then the content of his desire can be given by the declarative sentence 'I become a computer hobbyist'. Likewise for wishes, hopes, desires, fears, intentions, emotions with propositionally articulated objects, and so on. This wide definition of 'judgement' serves a theoretical purpose. Contemporary moral philosophy includes conflicting accounts of morality, each identifying moral judgements with states of different kinds, including beliefs,⁴ desires, and propositionally articulated emotions. Common sense parlance is similarly flexible, with 'judgement' sometimes bearing no cognitive implications beyond the equally neutral 'decision'. There is no reason to deny that states exist corresponding to each of these different uses, even if they are not identifiable with moral judgements as here understood. To say that a state is a judgement is therefore not to say anything of great philosophical significance. We can allow that some non-cognitive states express judgements.

Cognitivism as here understood is neutral on a number of controversies in moral psychology. In particular, it is neutral on whether human agency is fully explicable in terms of the causal powers of teleologically related beliefs and desires as postulated by a Humean theory of motivation.⁴ On this view, moral judgements are cognitive just in case they are Humean beliefs. The Humean account faces three

³ This hypothesis is neutral with respect to the existence of a complementary conception of cognitivism along primarily epistemological lines. On such a conception, cognitive states would be states which qualify as knowledge-apt, where the concept of knowledge is defined without reference to a substantial notion of representation (Cf. Wiggins (1990)).

⁴ Cf. Smith (1994).

difficulties avoided by the present approach. First, some states are not classifiable as Humean beliefs but have still been thought to exhibit enough marks of cognitive status to have correctness-conditions, even though the possession of correctness-conditions is one of the distinguishing marks of Humean beliefs.⁵ Second, some states, such as intentions and emotions, appear to play central roles in motivation without being classifiable as Humean beliefs, desires, or constructions thereof. Third, it is unclear how the Humean can account for the subdoxastic states shared by humans and lower animals, and which play some role in human motivation.

Cognitivism as here understood avoids these difficulties by defining cognitive states without reference to the Humean theory of motivation. First, if Humean non-beliefs exhibit independently defined marks of cognitive status, this will explain why they can be assigned correctness-conditions. Second, if states other than Humean beliefs and desires play a role in motivation, the present account gives a classification of cognitive status on which it could turn out that some non-Humean states are cognitive. Whatever the plausibility of the Humean theory of motivation, it is a virtue of cognitivism as here understood that it is neutral on these issues.

II. Marks of Cognitive Status

In what follows I examine the claim that some moral judgements have assertoric contents, are non-permissibly world-guided, and are motivationally detached. I shall not argue that all three marks are individually necessary for cognitive status. For example, it is controversial whether motivational detachment is a necessary condition for cognitive status. There are decent arguments indicating that some cognitive states are partially definable by their motivational efficacy.⁶ Moreover, the most developed arguments of this kind claim that some

5 Cf. Blackburn (1998b).

6 Cf. Millikan (1996).

moral judgements are motivationally efficacious. I shall not fully evaluate these arguments here, and so will not argue that there are no cognitive moral judgements possessing the Janus-like nature which led Altham to label them 'besires'.⁷ What I shall argue is that some moral judgements do not possess this feature, and are hence motivationally detached. Motivational detachment, even if not constitutive of moral judgements, is a mark of some moral judgements.

Assertoric Content

Assertoric content is a mark of cognitive status. The set of thoughts and utterances we commonly refer to as 'claims' all have as their paradigmatic form of linguistic expression declarative sentences standing in inferential relations, the natural logical systematisation of which is a truth-conditional semantics. Suppose Pepe thinks that either Manolo will write by snailmail or he is a computer hobbyist. But Manolo wrote to Pepe by email. From which Pepe concludes that Manolo is a computer hobbyist. How do we account for the soundness of Pepe's inference? The natural thing to say is that it forms a logically valid pattern, according to which an inference is valid if and only if it is impossible for all the premises to be true and the conclusion false. Pepe's inference exhibits this pattern as the first premise is a two-termed disjunction, the second premise denies the first disjunct, and the conclusion affirms the second disjunct. If the validity of Pepe's inference can be accounted for in this way, its constituent claims must be individually truth-apt. Only thus is it possible for premises and conclusion to be true together. The inferential articulation of judgements in truth-conditional patterns of inference is partially constitutive of those judgements exhibiting what Wright has called *assertoric content*.⁸

Following Wright, I shall take the basic marks of assertoric content to include the following four, each of which is explicable with reference to

7 Cf. Altham (1986).

8 Cf. Wright (1992).

the example above. First, assertoric contents have significant negations. We can infer from the fact that Manolo did not write to Pepe by snailmail that he is a computer hobbyist, given the disjunctive claim 'Manolo will write by snailmail or he is a computer hobbyist'. Second, assertoric contents can be embedded within conditionals. We can validly infer from the premises 'If Manolo is a computer hobbyist, then he will write by email', and 'Manolo is a computer hobbyist' the claim 'Manolo will write by email'. Third, assertoric contents can be embedded within propositional attitude ascriptions. Not only can we say that Pepe was furious with Manolo because he believed Manolo was a computer hobbyist. We can also say that insofar as Pepe is rational and believes both that if Manolo is a computer hobbyist then he will write by email, and that Manolo is a computer hobbyist, then Pepe will believe that Manolo will write by email. Fourth, assertoric contents have associated conditions of warranted belief and assertion, where these warrants are inheritable with warrant from one thinker to another. If Pepe is warranted in his belief that Manolo is a computer hobbyist, and Manolo is warranted in believing that Pepe is warranted in believing that Manolo is a computer hobbyist, then Manolo is warranted in believing that Manolo is a computer hobbyist. Inheritance of warrant can break down in at least two ways. First, Manolo may have no warrant to believe that Pepe's belief is warranted. He might have evidence to the contrary, such as knowledge that Pepe is prone to illegitimately infer that anyone who writes him by email is a computer hobbyist, or that Pepe's belief is false. Second, and regardless of Manolo's warrant, Pepe's belief might not be warranted. Perhaps Manolo was strongly encouraged to write by email by a trustworthy friend, and Pepe has no good evidence to suggest that Manolo is a computer hobbyist.

I shall assume with Wright that satisfaction of the basic marks for assertoric content qualifies a set of claims as truth-apt, by which I mean that they can be interpreted as *purporting* to be true, not that they are true. I shall also assume that the concept of truth so understood is

deflationary.⁹ In other words, the application of a truth-predicate entails no more than the existence of a disciplined meta-linguistic practice of attributing truth to judgements as a way of approving those judgements, whether this is a matter of expressing one's personal sympathy with them or measuring them against an independent reality. On this view, the concept of truth has no metaphysical import. The metaphysics of assertoric judgements is determined by considerations beyond their assertoricity alone.

The possession of assertoric content is arguably neither a necessary nor sufficient condition for the capacity to represent *per se*. First, a representational state may fail to possess assertoric contents if it is sufficiently primitive. A representational system lacking both language and self-consciousness may still be in states with representational content even if it cannot represent those states to itself, and is thereby unable to relate its own representations inferentially to each other and attribute truth and warrant to them. Some of the lower animals may be in this position with respect to their internal representations. The same could be true of artificial representational systems, such as signal boxes and thermostats. I shall not deny that such systems represent, even if assertoric contents cannot be literally attributed to them. Nor shall I deny that such systems can be described *as if* they have assertoric contents by representational systems whose contents are so articulated. It may well be a necessary condition of any representational system that its contents be thus describable. While not necessary for the ability to represent *per se*, assertoric content is a necessary condition for the kind of representational content associated with propositional knowledge in

9 Cf. Wright (1992), Horgan and Timmons (2001). Both assumptions can be denied. First, if Wright's marks of assertoric content do not entail truth-aptness, then the fact that moral judgements exhibit these marks does not show that they are truth-apt. I shall simply assume that moral judgements exhibit whatever marks are required for truth-aptness. Second, if truth-aptness entails a metaphysically substantial notion of representation, then the truth-aptness of moral judgements shows that they are cognitive as here understood. I shall grant the falsity of this claim to the non-cognitivist for the sake of argument. Cf. Blackburn (1999).

humans. This is the kind of representational purport attributed to moral judgements by moral cognitivism as here understood.

Assertoric content is not sufficient for representational purport. It is possible to understand acts of imagining, for example, as involving assertoric contents. Yet it is implausible to suppose that all imaginings purport to adequately represent the nature of an imaginary reality. Furthermore, it is a debatable whether possession of assertoric content qualifies moral judgements for ~~representational purport~~. Some non-cognitivist argue that the assertoric features of moral judgements are superficial, or merely linguistic, properties obscuring the underlying non-cognitive status of these judgements.¹⁰ On this view, there is more to being a cognitive state than sharing the characteristic linguistic articulation of a cognitive state. Instead, cognitive status is determined in part by the psychological nature of the mental states possessing assertoric contents, where this nature is determined by their underlying teleology. Cognitivism as here understood is consistent with the claim that the cognitive status of mental states is determined independently of their characteristic linguistic expression.

Moral judgements have assertoric contents, as given by the four marks of negation embedding, conditional embedding, propositional attitude embedding, and warrant acquisition. If Manolo believes that if Pepe will not accept an email approach, he ought to write by snailmail; and he believes that Pepe will not accept an email approach; then he is logically entitled to infer that he ought to write by snailmail. Furthermore, if Manolo is warranted in believing the premises of this inference, he is also warranted in believing the conclusion. And Pepe, who may or may not be warranted in believing Manolo's premises, will be sure to have the requisite collateral beliefs to inherit from Manolo his warrant in believing that he ought to write by snailmail. In terms of their inferential relations, moral judgements behave as paradigm assertoric judgements.

¹⁰ Cf. Blackburn (1998a, 1998b).

The above example brings out two debated features of moral judgements, each supporting the claim that they have assertoric contents. First, the contents of moral judgements occur as antecedents in conditionals in which they are not asserted. In these contexts, no attitude is expressed towards the contents themselves. When Manolo judges that if it is wrong to write to Pepe by email then he should write by snailmail, he is not thereby committed to the claim that it is wrong to write to Pepe by email. Nor does he, merely by making this judgement, express any attitude of approval or disapproval towards the claim that it is wrong to write to Pepe by email. Nevertheless, this content can play a significant role in a pattern of sound inference on the assumption that it is true-apt. Manolo can validly infer from the conjunction of this judgement and the further judgement that it is wrong to write to Pepe by snailmail, that it is not wrong to write to Pepe by email.

The second feature of moral judgements brought out by the above example is that their contents figure as premises and conclusions in arguments the premises of which are not all moral. Manolo can infer that he ought to write to Pepe by snailmail from a conditional premise which states a fact about Pepe's receptiveness to electronic approaches. Given that Manolo's inference is logically valid, and that we understand logical validity in terms of truth preservation, it follows that both the moral and non-moral claims making up the inference are truth-apt, and truth-apt in the same way.

The embedding of the contents of moral judgements in conditionals and other patterns of inference calling for a truth-conditional semantics has historically been considered a problem for non-cognitivism.¹¹ The present account allows that non-cognitive states without metaphysically substantial correctness conditions can be given a truth-apt linguistic articulation. Some non-cognitivists appear to hold exactly this view of the relationship between the semantics of moral judgements and the mental

¹¹ Cf. Geach (1964), Blackburn (1984, 1993), Hale (1986).

states they express.¹² Cognitivism as here understood differs from this hybrid version of non-cognitivism by claiming that moral judgements possess further, metaphysically substantial, marks of cognitive status.

World-Guidedness

It is common to distinguish mental states by their functional role. Talk of the functional role of mental states is sometimes made explicit in terms of the metaphor of constitutive aims. Thus, whereas the constitutive aim of belief is to correctly represent the environment, the constitutive aim of desire is to regulate that environment. The constitutive aims of different mental states are sometimes distinguished by the functional relationship between their contents and what these contents represent. Thus, whereas it is the constitutive function of beliefs to have their contents regulated by the nature of the environment, it is the constitutive function of desires to regulate the nature of the environment in accordance with their content. Beliefs aim to have their content fit the world. Desires aim to have the world fit their content. This difference in functional role underlies the distinction between beliefs and desires in terms of the metaphor of *directions of fit*.¹³ The direction of fit of beliefs is *world-guided*. The direction of fit of desires is *world-guiding*. As is manifest from the fact that beliefs are paradigmatic cognitive judgements, being world-guided is one mark of cognitive status.

I shall assume that being world-guided is a property of mental states themselves, in principle distinguishable from their characteristic form of linguistic expression. On this view, the correctness-conditions of world-guided mental states give independent metaphysical substance to the truth conditions of the linguistic claims expressing their contents. While world-guided mental states purport to accurately represent their domain, they may fail to do so in two ways. First, they may misrepresent facts internal to that domain, as when Pepe takes Manolo for a computer

¹² Cf. Blackburn (1998b).

¹³ Cf. Williams (1970), Humberstone (1992), Smith (1994), Velleman (2000). For doubts about the literal significance this metaphor, see Sobel and Copp (2001).

hobbyist. Second, they may falsely assign content to an empty domain, as when Pepe worries that humans are victims of an intangible cyberspace conspiracy.

World-guidedness is a necessary condition for cognitive status. No state can represent the world correctly or incorrectly unless it has correctness conditions given by how its content matches up with the features of the domain it purports to represent. In this sense, paradigmatic desires do not represent, as it is no functional defect of a paradigmatic desire that its content fails to correctly represent the agent's environment. The paradigmatic function of desires is to make the agent's environment fit their content, thereby tending to make useless or idle desires the contents of which correctly represent that environment (desires to maintain the status quo being an exception).

While being world-guided might be a sufficient condition for not being a desire, it is not sufficient for cognitive status. World-guidedness is not all or nothing.¹⁴ Most acts of imagining are world-guided, in that they aim to capture possibilities in ways constrained by determinate requirements on their imaginative construction. It is not enough in order to imagine a whale in the context of a zoology lesson to imagine a whale in just any manner. The whale should be considered in a certain environment, as possessing specific natural features, and so on. Nevertheless, the act of imagining leaves room for choice in the representation of the whale's features. There is no requirement that the object imagined corresponds to a determinate whale in the actual world. Likewise, when constructing a fiction, constraints of style and content are employed to make the fiction intelligible as a representation of entities of a certain kind. Hemingway's short story 'Hills like White Elephants' does not lend itself to depiction as an adventure among the Swiss Alps. Nevertheless, there is no requirement that the story be understood as a true description of a historical event. Fictions and imaginings exhibit a form of world-guidedness, although there is no requirement that either

¹⁴ Cf. Velleman (2000).

be literally correct. Fictions and imaginings are world-guidedly *permissivae*, and thereby fail to have correctness conditions, even though each can be undertaken well or badly within the constraints imposed by the relevant genre of invention. Nor are fictions or imaginings essentially world-guiding. Although images and fictions can be used to change the world, it is not obviously of their nature that they are so used. In this sense, imaginings and fictions are different from non-cognitive states like desires, which are essentially world-guiding insofar as their function is to make the world fit their content. Other apparently non-cognitive states might be thought to be both world-guided and world-guiding. Emotions such as fear may function to motivate agents to confront perceived danger, but can also misfire when the perceived danger does not exist.¹⁵ If emotions of fear include a commitment to the existence of danger, they are at least permissibly world-guided.

I shall assume that a state has correctness conditions only if it is non-permissibly world-guided. Such states only perform their function to the extent that they correctly reflect the facts in their domain. A belief is incorrect if any part of its content represents the world incorrectly (incomplete representation does not entail incorrect representation). It is to that extent defective qua belief. This is not to deny that the belief is marvellous in other ways. Complete representational adequacy is generally impracticable, and even pointless. A non-permissibly world-guided state is one whose functional role is at least partly defined by the aim to have a content correctly representing its constitutive domain. Not all such states may have representational purport. Signal boxes, thermostats, and the internal representations of lower animals may all be non-permissibly world-guided. They do not have representational purport unless they also have fully articulated assertoric contents, which they may not have.

It is arguable that mental states can be both world-guided and world-guiding. This possibility is overlooked if it is supposed that both

15 Cf. Oakley (1992).

directions of fit must attach to the same content. Suppose the belief that it is wrong to write to Pepe by email has the direction of fit of both belief and desire. It might then seem that it would aim to both come into existence and go out of existence if its content is true – its nature as belief suggesting the former, its nature as desire suggesting the latter. However, a mental state could have the directions of fit of both belief and desire provided the directions of fit are associated with distinct contents.¹⁶ Suppose Manolo judges it wrong to write to Pepe by email. Qua belief, Manolo's mental state will aim to stay in existence as long as it is warranted. In the meantime, Manolo's judgement can function to change the world. There is an indefinite number of contents Manolo could have a disposition to make true in virtue of his judgement that it is wrong to write to Pepe by email. He might desire to write by snailmail, to erase Pepe from his address book, or to pursue some other Pepe-related end. The desiring of at least one such content to some arbitrary degree is all that is necessary for Manolo's judgement being world-guiding.¹⁷ A mental state being world-guiding is therefore arguably consistent with its being non-permissibly world-guided, and therefore also with its being a judgement with representational purport.

This possibly dual function of world-guided states illuminates another controversy. It is sometimes suggested that moral judgements cannot be cognitive because their underlying teleology is non-representational, as given by a functional account of their existence as naturally evolved coordinating devices.¹⁸ This argument is unconvincing if it is assumed that a mental state can have either a world-guided (representational) or a world-guiding (co-ordinating) teleology, but not both. A mental state can have both a world-guided and world-guiding teleology if the world-guiding function of that state explains the world-guided function of that

16 Cf. Smith (1994).

17 Arbitrary, because *pace* Lewis there is no *a priori* requirement that strength of desire track strength of belief, except possibly in agents who are ideally rational in decision-theoretic terms. Cf. Lewis (1996).

18 Cf. Ruse (1986), Gibbard (1990).

state. The co-ordinating function of moral judgements would explain the representational purport of moral judgements if their representational purport evolved in aid of their co-ordinating function. The claim that moral judgements are non-permissively world-guided is consistent with the claim that they have an underlying world-guiding teleology. One way for mental states to co-ordinate could be by including moral judgements aiming to be true.

Moral judgements exhibit obvious marks of non-permissive world-guidedness. Moral thinking works on the assumption that moral judgements are responsive to reasons the nature and weight of which provide differential warrant for judgements.¹⁹ When someone makes a moral judgement with which we disagree, we look for reasons to sway that person to change his or her judgement. We regard agents who fail to be swayed by reasons we find conclusive as defective in their judgement. Agents who treat moral problems as merely imaginative exercises are not regarded as engaging in genuine moral thought. The function of moral judgements is in part to have their content match the content of moral reasons. Insofar as moral judgements function badly if their contents fail to track moral reasons, moral judgements are non-permissively world-guided. It is compatible with this claim that reasons can run out. There may be no grounds on which to decide whether email is better than snailmail. It does not follow that the claim 'Email is better than snailmail' cannot be non-permissively world-guided. First, there could be no fact about which of email and snailmail is better, even though they both surpass the use of carrier pigeon. Second, the two methods of communication could be equally good. If they are, the fact is that there is good reason to choose one or the other and not as much reason to choose carrier pigeon.²⁰

19 By 'reason' I mean 'consideration speaking for or against a given set of options'. This use is compatible with different views about what kinds of consideration are morally basic; be they goods, values, rights, virtues, and so on.

20 Cf. Brink (1994)

Moral judgements would be no more than permissively world-guided if their function were to correctly represent non-moral facts, leaving the linking of non-moral and moral features to the discretion of the individual. There is a respectable history of arguments for non-cognitivism based on this and similar ideas.²¹ Yet the basic norms of moral reflection rule out any such discretion as constitutive of moral judgements. Moral judgements are neither more nor less vulnerable to error with respect to the linking of non-moral and moral features than they are with respect to links between moral features themselves.²² Pepe's insistence on snailmail on the grounds that his email was installed on a Tuesday is as ridiculous as the claim that cowardice is a greater virtue than courage.

The non-permissive world-guidedness of moral judgements entails that they have correctness conditions definable independently of the deflationary concept of truth entailed by their assertoricity alone. It follows that the correctness conditions of moral judgements are not fully explicable in terms of a concept of truth employed in meta-linguistic speech acts infused with moral approval, but devoid of metaphysical commitment.²³ The non-permissive world-guidedness of moral judgements is powerful evidence in favour of cognitivism.

The fact that moral judgements *appear* to be non-permissively world-guided is no proof that they really *are* non-permissively world-guided. Perhaps the assertoricity of moral judgements functions as a cunning disguise for their uniquely world-guiding nature. True, the surface features of common sense morality are suggestive of the existence of metaphysically substantial correctness conditions for moral judgements.²⁴ But this could be a misleading feature of the way in which the assertoric contents of uniquely world-guiding states are inferentially articulated. Taking this line, non-cognitivists such as Blackburn have argued that

21 Cf. Stevenson (1944), Hare (1952), Gibbard (1990).

22 Cf. Blackburn (1984, 1993, 1998b).

23 Cf. Horgan and Timmons (2001).

24 Cf. Lillehammer (1999, 2000).

claims such as 'Murder is wrong no matter what anyone would think about it' should be interpreted as the endorsement of moral norms rejecting the attitude of letting what people think settle moral questions.²⁵ On this view, claims apparently about metaphysically substantial truth conditions are more fruitfully interpreted as pieces of moral advice, not as meta-ethical claims about the relationship between moral judgements and their correctness conditions. If so, the apparently non-permissive world-guidedness of moral judgements could be a misleading effect of their metaphysically neutral assertoricity.

Fruitful advice to the contrary, the cognitivist holds the higher ground in this dispute. First, the cognitivist has an obvious explanation why it is good advice not to let what people think settle moral questions, namely that moral facts are not determined by what people think. Second, while there is at least some good evidence that moral judgements are non-permissively world guided, there is no good evidence that moral judgements are uniquely world-guiding. True, it is sometimes argued that moral judgements are uniquely world-guiding because they display the motivating efficacy of affective states like desires.²⁶ But this argument is unsound. First, moral judgements can be both world-guided and world-guiding if their being world-guiding explains their being world-guided and their different directions of fit attach to distinct contents. Second, moral judgements do not all display the motivating efficacy of affective states like desire. In fact, some moral judgements display the motivational detachment of cognitive states like belief.

Motivational Detachment

If assertoric content and non-permissive world-guidedness are jointly sufficient to make moral judgements cognitive, why worry about motivational detachment? We have seen that motivational detachment is

²⁵ Cf. Blackburn (1993).

²⁶ Cf. Hare (1963), Blackburn (1984).

not a necessary condition of cognitive status if there are *désires*. Furthermore, imaginings are as motivationally detached as beliefs, but are at best permissibly world-guided. Higher-order desires are not world-guided, but have no motivational implications for action (although they do have motivational implications for the desires they take as their objects). Manolo's idle wish that computers had never been invented is no more world-guided than his desire that they be all destroyed. Yet in contrast to paradigm desires it is not obvious that idle wishes exhibit more motivational efficacy than imaginings (although some motivational efficacy would explain the difference between wishing that P and merely entertaining that P). Either way, motivational detachment is not a sufficient condition for cognitive status. So why the centrality of motivational detachment in the debate about cognitivism?²⁷

One explanation would be excessive attention to the possibility that all non-cognitive states are motivationally engaged. Idle wishes aside, this could be true, but it does not follow that all motivationally engaged states are non-cognitive. Another explanation is the centrality of the Humean theory of motivation to the debate, beliefs being paradigm examples of motivational detachment and desires being paradigm cases of motivational engagement. Excessive attention to the Humean theory could lead to the view that motivational detachment equals belief, while motivational engagement equals desire. While mistaken, this view points to another possibility. The set of motivationally detached and non-permissively world-guided states with assertoric contents could be co-extensive with the set of Humean beliefs. Not only is this arguably true, it is arguable that some moral judgements belong to this class.

A mental state is motivationally detached if there is no interesting conceptual connection between being in that state and being motivated in some way. Humean beliefs are paradigms of the motivationally detached. If Manolo believes that Pepe hates email, nothing follows about Manolo's motives. It all depends on what Manolo feels about such

²⁷ Cf. Brink (1989), Dancy (1993), Smith (1994).

things as emails, the thwarting of other people's desires, Pepe's feelings, and so on. On the other hand, once we learn that Manolo desires to avoid any inconvenience to Pepe and so on, we can infer how Manolo will be motivated, even if we cannot predict how he will act. Humean desires are paradigms of the motivationally engaged.

Some moral judgements are motivationally detached in the way that Humean beliefs are. Their motivational detachment can be illustrated by comparing a range of successively less extreme cases of moral disengagement. The most extreme case is a familiar philosophical stock character, namely the amoralist. For the amoralist, moral considerations are said to have no motivational force whatever.²⁸ The amoralist is not weak-willed or depressed, nor does he fail to desire ends he desires to desire, or desire ends he desires not to desire. The amoralist is indifferent or hostile to what he judges morality to require. He may be perfectly knowledgeable about the requirements of morality. His intimate knowledge of the nuances of right and wrong might render him a perfectly competent advisor in moral matters. At the same time, the amoralist has no inclination to promote the ends which morality embodies, either because these ends are incompatible with ends he cares about or because he does not care about very much at all. Pepe might know that the moral requirement to not upset Manolo's feelings is based on the altruistic assumption that the feelings of others count as much as one's own. Pepe, however, might be so selfish that his concern for others is merely instrumental to the pursuit of his own ends, none of which will be thwarted by his accusing Manolo of being a computer hobbyist. He might therefore not care if it is wrong for him to upset Manolo's feelings by doing so.

Complete amoralists are hard to come by, perhaps even impossible or incoherent.²⁹ One stock example in the philosophical literature is the complete egoist, who fits the bill perfectly on the assumption that

28 Cf. Brink (1989), Smith (1994).

29 Cf. Hare (1963), Smith (1994).

morality is essentially other-regarding. Another stock example is the psychopath, who is often thought of as someone who, even if clear about the rights and wrongs of the case, is unwilling or unable to affectively take up another person's viewpoint.³⁰ Whether psychopaths are as factually lucid or affectively challenged as their philosophical caricature suggests is not a question to settle here. Whatever our favourite example of amorality, we can distinguish between an amoralist who is unable to be moved by moral considerations, and an amoralist who is perfectly able, but simply does not care enough to be moved by moral considerations. We can furthermore distinguish between an amoralist who is willingly unmoved by moral considerations permanently and an amoralist who is willingly unmoved by moral considerations only temporarily. Finally, we can distinguish between an amoralist who is willingly unmoved by all moral considerations temporarily and an amoralist who is willingly unmoved by a subset of moral considerations temporarily. The latter specimen is not so hard to find, and may be exemplified by most of us part of the time. Manolo might lose his interest in his job for a period, during which he neglects moral requirements binding him to his colleagues. Pepe might be unconcerned for the good of anyone who fails to care about him and his career. Selective amorality is a much more plausible feature of real humans than pure egoism or psychopathy. It is also a common source of deep worries motivating scepticism about morality's demands. The feeling that promotion of an impartial good is at odds with personal convenience and the thought that morality threatens to undermine one's personal integrity are not always easily distinguishable.

If agents can be selectively amoral, there can be areas of moral importance to which they are indifferent, such as the good of colleagues, strangers, distant humans, sentient non-human beings, and so on. There can also be times at which morality fades into practical insignificance, such as periods of intense selfishness associated with youth, amid

30 Cf. Gaus (1990).

trappings of wealth, or conditions of extreme scarcity. Corresponding to the domain of moral concern and its associated time-span we can identify degrees of amorality. Even if specimens of the highest degrees of amorality are impossible, selective amorality is both possible and common. The possibility of selective amorality is evidence for the motivational detachment of the moral judgements to which selective amorality is indifferent or hostile. If agents can be selectively amoral, then some moral judgements are as motivationally detached as Humean beliefs are.

The construal of moral judgements as motivationally detached carries with it a burden of explanation. Some uses of the term 'moral judgement' refer to motivationally engaged states. It is platitudinous that a man's principles are revealed in his actions and that an agent's emotional state often tracks her moral judgement. Moral judgements are intimately connected to emotional responses such as guilt, shame, resentment and anger. In fact, these emotions are plausibly classifiable as distinctively moral. After all, what are morally shameful actions if not actions for which the feeling of shame is appropriate?

The construal of moral judgements as motivationally detached is compatible with these platitudes. First, the fact that certain cognitive judgements are motivationally detached does not entail that all judgements are. The present account is compatible with complementary senses of 'judgement', where the mental states picked out by the term are not motivationally detached. There might be a sense of 'judgement' according to which an agent's judgement can be read off from her actions, as when Manolo says that Pepe is constantly let down by his judgement. There may also be a sense of 'judgement' according to which an agent's judgement entails the existence of an intention or desire to comply with it. This would clearly be true of desires. Furthermore, this sense of 'judgement' is central to the idea that moral agents act in the light of their endorsement of norms. Moral education involves the internalisation of general moral guidelines, such as prohibitions against lying and theft, and requirements of generosity and truthfulness. The

endorsement of a moral norm may require that one either intends or desires to conform to it. If we understand making a moral judgement as involving the endorsement of norms, then moral judgements are not motivationally detached.³¹

It is possible to endorse a moral norm without having critically reflected on it. Children adopt moral principles before they learn to justify them. Furthermore, the primitive practices from which contemporary morality developed may have had little or no place for motivationally detached moral judgements aiming to determine the grounds of moral norms. Even so, human history has produced increasingly sophisticated forms of reflection on the grounds of moral norms. Whatever its motivationally engaged origins, the contemporary moral mind is capable of abstract reflection on the grounds of moral norms. This capacity for abstract reflection has brought arbitrarily high degrees of motivational detachment in its wake. As the case of the selective amoralist illustrates, you can make a judgement on the grounds of a moral norm and fail to be accordingly motivated. You might judge that the principle of truthfulness towards which you have been committed is groundless, yet continue to intend to observe it. Alternatively, you might judge that it is a morally sound principle, yet cease to intend conforming to it because you have lost your interest in the ends promoted by truthfulness. Either way, your moral judgement can be motivationally detached.

Similar conclusions can be drawn for the connection between moral judgements and moral emotions, such as shame, guilt, resentment, and anger. There is a sense of 'moral judgement' in which an agent's judgement can be read off from how he feels about a given situation. A state of rage at someone's intrusiveness is classifiable as a negative moral judgement, at least if the state of rage is accompanied by a positive attitude towards one's rage. A feeling of guilt at having hurt someone, when accompanied by a further desire to preserve this feeling until one

31 Cf. Copp (1995).

has asked forgiveness, is classifiable as a negative judgement about what one has done. Even a simple feeling of joy or disgust could be so classifiable, as instanced by the crude examples of approval or disapproval appealed to by the logical positivists in their initial formulations of emotivism.³² As with states of moral endorsement, the existence of such emotional states is fundamental to moral education. Higher-order feelings of approval of emotional reactions are not fully motivationally detached. Simple feelings of joy or disgust are even less so. Nevertheless, the existence of such states does not undermine the claim that there is a further set of states, associated with the reasoned assessment of emotions, which are motivationally detached. Furthermore, the fact that moral judgements are motivationally detached does not exclude them from causing intentions, desires, and emotions, which then provide motivation to act accordingly when combined with suitable beliefs about the means of doing so. The relation between motivationally detached moral judgements and such motivationally engaged states is contingent, causal, and defeasible.³³

III. Concluding Remarks

I have defended the claim that moral judgements have representational purport, and that they are in that sense cognitive. Representational purport is a metaphysically neutral notion. A state may have representational purport even if the domain it purports to represent is empty. Just as theories and claims can be false because the entities they postulate fail to exist, a whole system of representations may fail in its representational purport because the entities by which it purports to be guided do not exist. At different times, error theories of different kinds have been defended about such diverse topics as the theoretical postulates of physical science, medium-sized dry goods, causes, colours,

³² Cf. Ayer (1936).

³³ While this claim is consistent with the view that moral judgements cause relevant motives on pains of irrationality, it does not entail it. Cf. Smith (1994), Lillehammer (1997), and Svavarsdottir (2000).

the self, religion, astrology, abstract objects, modality, morality, and value. The notion of representational purport provides one currency in terms of which different cases of possible error can be evaluated on their own terms, consistently with the assumption that the judgements in question aim to represent a genuine domain of fact.³⁴

University of Cambridge

References

- Altham, J. E. J. 'The Legacy of Emotivism' in G. MacDonald and C. Wright (eds) *Fact, Science, and Morality* (Oxford: Basil Blackwell, 1986).
- Ayer, A. J. *Language, Truth and Logic* (London: Victor Gollancz, 1936).
- Blackburn, S. *Essays in Quasi-Realism* (Oxford: Oxford University Press, 1993).
- 'Wittgenstein, Wright, Rorty, and Minimalism' *Mind* 107 (1998a), 157-181.
- *Ruling Passions* (Oxford: Oxford University Press, 1998b).
- Brink, D. O. *Moral Realism and the Foundations of Ethics* (Cambridge: Cambridge University Press, 1989).
- 'Moral Conflict and its Structure' in E. Mason (ed.) *Moral Dilemmas and Moral Theory* (Oxford: Oxford University Press, 1996), 102-126.
- Copp, D. *Morality, Normativity, and Society* (Oxford: Oxford University Press, 1995).
- Dancy, J. *Moral Reasons* (Oxford: Basil Blackwell, 1993).
- Gaus, G. F. *Value and Justification* (Cambridge: Cambridge University Press, 1990).
- Geach, P. 'Assertion' *Philosophical Review* 74 (1964).
- Gibbard, A. *Wise Choices, Apt Feelings* (Oxford: Oxford University Press, 1990).

³⁴ I thank the Editor and two anonymous referees for valuable comments.

- Hale, B. 'The Compleat Projectivist' *Philosophical Quarterly* 36 (1986), 65-84.
- Hare, R. M. *The Language of Morals* (Oxford: Oxford University Press, 1952).
- *Freedom and Reason* (Oxford: Oxford University Press, 1963).
- Horgan, T. and Timmons, M. 'Nondescriptivist Cognitivism: Framework for a New Metaethical' *Philosophical Papers* 29 (2000), 121-53.
- Humberstone, L. 'Direction of Fit' *Mind* 101 (1992), 59-83.
- Johnston, M. 'Objectivity Refigured' in J. Haldane and C. Wright (eds) *Reality Representation, and Projection* (Oxford: Oxford University Press, 1993), 85-130.
- Lewis, D. 'Desire as Belief II' *Mind* 105 (1996), 303-314.
- Lillehammer, H. 'Smith on Moral Fetishism' *Analysis* 57 (1997), 187-195.
- 'Analytical Dispositionalism and Practical Reason' *Ethical Theory and Moral Practice* 2 (1999), 117-133.
- 'Revisionary Dispositionalism and Practical Reason' *Journal of Ethics* 4 (2000), 173-190.
- Millikan R. G. 'Pushmi-Pullyou Representations' in L. May, M. Friedman, and A. Clark (eds) *Mind and Morals* (Cambridge: MIT Press, 1996).
- Oakley, J. *Morality and the Emotions* (London: Routledge, 1992).
- Pink, T. *The Psychology of Freedom* (Cambridge: Cambridge University Press, 1991).
- Ruse, M. *Taking Darwin Seriously* (Oxford: Basil Blackwell, 1986).
- Smith, M. *The Moral Problem* (Oxford: Basil Blackwell, 1994).
- Sobel, H. and Copp, D. 'Against Direction of Fit Accounts of Belief and Desire' *Analysis* 61 (2001), 44-53.
- Stevenson, C. L. *Ethics and Language* (New Haven: Yale University Press, 1944).
- Svavarsdottir, S. 'Moral Cognitivism and Motivation' *Philosophical Review* 108 (2000).
- Velleman, D. J. 'On the Aim of Belief' in Velleman D. J. *The Possibility of Practical Reason* (Oxford: Oxford University Press, 2000).
- Wiggins, D. 'Moral Cognitivism, Moral Realism, and Motivating Moral Beliefs' *Proceedings of the Aristotelian Society* 91 (1990), 61-86.
- Williams, B. 'Deciding to Believe' in Williams, B. *Problems of the Self* (Cambridge: Cambridge University Press, 1970), 136-151.
- Wright, C. *Truth and Objectivity* (Cambridge: Harvard University Press, 1992).