

Moral Value

Psychology, Metaphysics, Semantics

Uriah Kriegel

the unemotional intellect may carry us into a mathematical dreamland where nothing is but what is not (George Eliot, Daniel Deronda)

Introduction/Abstract

This is a position paper. It presents a cohesive framework that addresses some of the defining issues of contemporary metaethics, notably the nature of moral judgment, moral reality, and moral language. The framework is supposed to appeal to philosophers antecedently attracted, on the one hand, to the idea that there are no mind-independent entities that are values, and on the other hand, to the idea that there is still such a thing as substantive moral truth. §1 introduces three prominent divides in contemporary metaethics: between cognitivism and noncognitivism in moral psychology, between moral realism and antirealism in moral metaphysics, and between descriptivism and expressivism in moral semantics. §2 then presents, telegraphically dogmatically, a comprehensive approach to the mind that I call impure intentionalism; it type-individuates mental states in terms of their *intentional character*, understood as a combination of content *and attitude*. Finally, applying impure intentionalism to moral psychology, §3 distinguishes between two kinds of moral judgment, one cognitive and one noncognitive, and crafts a moral psychology, a moral metaphysics, and a moral semantics around this distinction.

1 *Three Problems of Contemporary Metaethics*

Modern philosophy organizes much of its thinking around the triangle of mind, language, and reality. Contemporary metaethics follows this pattern, with much of the debate focusing on (i) the psychology of moral judgments, (ii) the metaphysics of moral properties, and (iii) the semantics of moral utterances. In each of these areas, there is a defining opposition between two major perspectives.

Moral Psychology: Cognitivism vs. Noncognitivism

According to cognitivism, moral judgments are cognitive, belief-like states. This is taken to imply, most centrally, that they are (i) truth-apt and (ii) not inherently motivational. That is, a moral judgment can be assessed for truth or falsehood and its impact on motivation and ultimately action is indirect in the sense that it requires the aid of other mental states to yield action. According to noncognitivism, now, moral judgments are noncognitive, desire-like states. This means that they are (i) not truth-apt but (ii) inherently motivating. In the background is the idea that mental states in general divide into two relatively neat groups, the truth-apt and extrinsically motivating versus the truth-inapt and intrinsically motivating, that do not tolerate unholy admixtures; something like this is sometimes referred to as the Humean theory of motivation (Smith 1987).

Moral Semantics: Descriptivism vs. Expressivism

According to descriptivism, moral statements such as 'genocide is wrong' are descriptive. This is taken to imply, most notably, that they express cognitive mental states and are therefore themselves truth-apt. Essentially, their indicative surface grammar is true to their nature and they integrate straightforwardly into inference patterns. According to expressivism, in contrast, moral statements express noncognitive states and hence are not truth-apt. Their indicative surface grammar is misleading and in truth their function is to express the speaker's personal feelings

about some object or event (Ayer 1936), to commend that object or event (Hare 1952), or to prescribe certain actions towards it (Stevenson 1944).

Moral Metaphysics: Moral Realism vs Antirealism

The terms 'realism' and 'antirealism' are fraught with multiple layers of ambiguity. Two notions of realism will interest us here, which I call *strong realism* and *weak realism*. Strong moral realism is the thesis that there are mind-independent moral facts. Weak moral realism is the thesis that there are moral facts, whether mind-dependent or -independent. I think of facts as states of affairs which obtain, and of states of affairs as either (a) some object's having a property or (b) some plurality of objects entertaining a relation. A fact is mind-dependent just if the property or relation involved in it is a *response-dependent* property or relation, that is, a property or relation whose instantiation conditions can only be specified by reference to a mental state of some subject's. Where moral antirealism denies that there are any moral states of affairs which obtain, weak moral realism claims that there are obtaining moral states of affairs, albeit ones involving response-dependent moral properties or relations, while strong moral realism insists that there are also obtaining states of affairs featuring response-independent moral properties as constituents.

2 The Impure Intentionalist Theory of Mind and the Nature of Attitudes

At a minimum, a philosophical theory of mind should provide us with a principle for type-individuating mental states. An intentionalist theory of mind is one that proposes to type-individuate mental states in terms of their intentionality. Pure intentionalism type-individuates them in terms of their intentional content; impure intentionalism does so in terms of their intentional *character*, where this is understood to involve both content and attitude. In other words, according to pure intentionalism, for any pair of type-different mental states M and M*, there is some difference in the contents of M and M*; according to impure intentionalism, M and M* may be type-different mental

states even if their contents are strictly the same, provided they employ different attitudes.

A full development of impure intentionalism would incorporate a theory of content and a theory of attitude. Here I will only lay out those aspects of my version of impure intentionalism that will be relevant to our treatment of moral judgment.

Content: Propositional and Objectual

The theory of content features a number of central issues, including: What if any are the naturalistic grounds of mental content? Does mental content supervene on subjects' nonrelational properties? What is the structure of Content? Is there nonconceptual content? Is there non-propositional content? For my purposes here, I want to remain neutral on all but the last of these questions. My only commitment in this area is this: Not all mental content is propositional.

It is clear that many psychological reports have a surface grammar that suggests non-propositional content: 'I see a cat,' 'You imagine a turtle,' 'He believes in ghosts,' 'She loves her son.' For at least some of these, this surface grammar may well be misleading. Thus, it is natural to think that 'He believes in ghosts' reports a mental state with the propositional content <there are ghosts>. To that extent, 'He believes in ghosts' can and should be paraphrased into 'He believes that there are ghosts.' My claim in this area, however, is that for some mental states, no propositional paraphrase is appropriate. When I say 'I love you' to my child, there is no F such that 'I love that you are F' communicates exactly what 'I love you' does, and more generally, there are no attitudes A_1, \dots, A_n and propositions p_1, \dots, p_n such that 'I A_1 that p_1 & ... & I A_n that p_n ' is an acceptable paraphrase of 'I love you.' We may say that 'I love you' expresses an irreducibly *objectual* attitude.

Attitude: The Representational-Guise Account

The attitudes are relations between subjects and contents: belief is a relation between a believer and what she believes, desire a relation between a desirous subject and what she desires, and so on. On my view, belief, desire, and our other attitudes are all

modifications of a single relation between subjects and contents: the intentionality relation. This is a relation in virtue of which a subject is *aware* of a content, *apprehends* a content, or *represents* a content. (These terms I use as different attempts to zoom in on the same phenomenon.) What distinguishes the attitudes from one another, on my view, is the guise under which they represent their contents. More precisely, each attitude *just is* a specific guise under which contents are represented by subjects. Thus, to *believe* that *p* is to represent *p* under the guise of the true ('*sub specie veri*'); to *desire* that *p* (or to desire *x*) is to represent *p* (or *x*) under the guise of the good ('*sub specie boni*'); and so on.

What are these representational guises? When a mental state represents *p* under the guise of the *F*, there is a sense in which the state casts *p* in an *F*-ish light. My desire that the restaurant have vegetarian options, for instance, casts the restaurant's having vegetarian options in a positive light; my belief that the sun is shining casts the proposition <the sun is shining> in 'a truth-y light,' and so on. At the same time, this kind of 'casting *p* in an *F*-ish light' remains external to the *content* of the mental state – it is not part of *what the state is about*. A belief that *p* may cast *p* as true, but its content is still <*p*>, not <*p* is true> – on pain of vicious regress if nothing else. Likewise, what the desire that the restaurant have vegetarian options is *about* is the restaurant's having vegetarian options, *not* it being good that the restaurant have vegetarian options. I capture this by saying that when a mental state represents *p* under the guise of the *F*, the state does not represent *p* as *F*, but instead *represent-as-F p*. Thus, a belief that *p* does not represent *p* as true, but instead *represents-as-true p*; a desire that *p* does not represent *p* as good but rather *represents-as-good p*. More generally, for each attitude there is a distinctive representational guise that characterizes it: fear of *x* represents-as-dangerous *x*, admiration of *y* represents-as-admirable *y*, and so on.

A note on the term 'desire' is desirable here, as the term leads a double life in contemporary philosophy of mind. In many discussions, the term is used rather technically, to cover a range of mental states that involve a favorable attitude toward their contents. These are the so-called pro attitudes: intention, wish, hope, decision, admiration, willingness, preference, trying, and more. Sometimes, though, 'desire' is used in a much narrower sense (essentially, the way it is used in everyday life). In this more specific sense, desire is *one among* the pro attitudes, alongside intention, wish, and the like. I point this out because I am making the claim that desire represents

under the guise of the good only for the wider notion of desire, and am making it with the most generic notion of good in mind (the *summum genus* of value, if you will). The claim is, really, that any pro attitude, in virtue of being a pro attitude, represents-as-good. However, each specific type of pro attitude deploys a more specific kind of representing-as-good. Thus, we noted above that admiration represents-as-admirable. Just as being admirable is a way of being good, representing-as-admirable is a way of representing-as-good. It follows that desire in the narrower, more mundane sense must exhibit its own specific way of representing-as-good. Perhaps it is representing-as-*achievably-good*, perhaps something else.

3 A Metaethical Framework

My metaethical picture is constrained by a number of starting-point desiderata. First and foremost among these is that there are no such mind-independent entities as values. My resistance to such entities is purely ontological and flows from a broadly nominalistic sensibility: I do not believe in values for the same reason I do not believe in numbers, in Platonic universals, and so on. The second desideratum is that, despite this, such statements as 'Genocide is wrong' and 'World peace is good' are literally and non-deflationarily *true* – and that this requires that they have *truthmakers*. A third important desideratum is that some intimate tie link moral judgment to moral motivation. There are other more minor desiderata, but these three are paramount. The first is essentially a demand for rejecting (what I called) strong moral realism, the second a demand for complying with descriptivism, and the third a demand to respect the basic insight of noncognitivism. The question for me is how to combine noncognitivism, descriptivism, and rejection of strong realism in one package deal. I will now present a framework that manages, I believe, to co-satisfy these desiderata. If you don't share these starting-point desiderata, mind you, you may find the exercise uncaptivating.

Moral Psychology: Two Kinds of Moral Judgment; Weak Noncognitivism

It is possible to be deathly afraid of flying without believing that flying is dangerous, and conversely to believe that flying is dangerous and yet find oneself unencumbered by fear of flying. This is because although both fear of flying and belief that flying is dangerous in some way cast flying as dangerous, they do so in different ways. The belief does so by explicitly attributing the property of being dangerous to flying, as part of its content; fear, in contrast, does so by representing flying under the guise of the dangerous. We may display the difference between the two intentional structures as follows:

danger belief ::	represent-as-true <flying is dangerous>
fear ::	represent-as-dangerous <flying>

The belief has a danger-invoking propositional content and represents that content under the guise of the true. The fear has a simple (danger-free) *objectual* content and represents that content under the guise of the dangerous. We see that the mind has two fundamentally different ways to commit to the dangerousness of flying – through content and through attitude. Why the mind has this two-track design, fielding both doxastic and emotive ways of committing to something's being dangerous, is an interesting question I will not get into here (but see Kriegel 2012).

My key thesis is that the same double track attends the mind's commitment to moral value. There are two kinds of moral judgment in our psychological repertoire, with thoroughly different intentional structures, which I call 'primary moral judgment' and 'secondary moral judgment'; we may display their respective intentional structures as follows:

secondary moral judgment ::	represent-as-true <world peace is morally good>
primary moral judgment ::	represent-as-morally-good <world peace>

Note that primary moral judgment is a specific kind of pro attitude: representing-as-morally-good is a way of representing-as-good. Secondary moral judgment is just a belief with a moral subject matter. Thus secondary moral judgment is a kind of belief, hence a *cognitive* state – truth-apt and merely extrinsically motivating – while primary moral judgment is a kind of pro attitude, hence a *noncognitive* state – truth-*inapt* and *intrinsically* motivating.

Two further clarifications are important. First, the intentional structures displayed above illustrate in truth only *positive* moral judgments. There is a similar distinction to make for *negative* moral judgments:

secondary moral judgment :: represent-as-true <genocide is morally bad>
primary moral judgment :: represent-as-morally-bad <genocide>

Secondly, for all I have said, primary moral judgment may either (a) be a specific, probably *sui generis* mental state, or (b) be a *species* of mental state, covering a number of different subspecies, each with its own distinctive way of representing-as-morally-good or representing-as-morally-bad. I lean toward (b), with the thought that moral emotions are the *de facto* primary moral judgments in our mental life. Certain kinds of respect, approval, appreciation and so on exhibit each its own manner of representing-as-morally-good, while certain kinds of indignation, contempt, guilt, and so on exhibit each its own manner of representing-as-morally-bad. A full development of this moral psychology will map out the variety of guises of the morally good/bad featured in our mental life.

The thesis that there are these two categorically different kinds of moral judgment, one cognitive and one noncognitive, does not fit well into standard forms of either cognitivism or noncognitivism. Nonetheless, I designate the noncognitive variety of moral judgment 'primary' for a reason. We will be in a position to appreciate this reason momentarily, but the idea is that the noncognitive variety of moral judgment is more fundamental than the cognitive variety in an important way. In consequence, the view is committed to the following thesis of recognizably noncognitivist bent: all *fundamental* moral judgments are noncognitive states. Call this *weak noncognitivism*.

Moral Metaphysics: Response-Dependent Truthmakers; Weak Moral Antirealism

Primary moral judgments are not truth-apt, but they are not the only kinds of moral judgment in our psychological repertoire. We also entertain moral beliefs, which *are* truth-apt. Are any of them actually true? A negative answer results in so-called error theory (Mackie 1977); a positive answer admits of moral truths.

I answer in the positive. And while some philosophers who admit of moral truth do so only in a minimalist, deflationary sense (e.g., Horgan & Timmons 2006), the moral truth I want in my metaethics is substantive and requires real-world truthmakers. However, since I do not believe in mind-independent entities that are values, these truthmakers cannot be facts consisting in world peace and genocide instantiating such values (or 'participating in' them in a Platonic vein). Instead, I take the truthmaker of a moral belief (i.e., *secondary* moral judgment) that x is morally good/bad to be the fact that x is disposed to elicit the right response in the right respondents (under the right conditions). The *right response* is the entertaining of a *primary* moral judgment about x (i.e., a noncognitive state with the objectual content <x> represented under the guise of the good/bad). The *right respondent* is a certain properly idealized version of us – where (i) 'us' in the present context covers any creature capable of primary moral judgment and (ii) 'properly idealized' is still to be unpacked (the extant literature offers a number of initial options). For instance, the truthmaker of the belief that genocide is morally bad is the fact that genocide is disposed to elicit indignation about it (or a similar negative moral emotion) in a properly idealized version of us.

Facts such as this are *response-dependent* facts, facts where something has a property whose instantiation conditions cannot be specified without reference to someone's mental state, namely, an idealized subject's primary moral judgment. These are the only facts populating moral reality, I contend, and are the facts our secondary (i.e., cognitive) moral judgments effectively track. In other words, my view is that there are no mind-independent moral facts, but there are *mind-dependent* moral facts, which constitute the truthmakers of our (true) moral beliefs. (To say that is not to say that the *phenomenology* of our moral experience is as of encounter with such response-dependent facts. And to say *this* is not to say that the phenomenology is as of encounter with response-*independent* facts. There is important work on the phenomenology of moral judgment yet to be done.)

This is essentially a form of weak moral realism. If it be objected that weak realism is a rubber duck (i.e., is not really a kind of realism), I would probably agree.

Are these response-dependent moral facts 'objective'? There are many different ways to understand this question. One useful way would be settled as follows: if *all* idealized versions of us would respond with the same type of primary moral judgment

to x (either positive or negative), then the fact that x is morally good or morally wrong is an *objective* fact; if some idealized versions of us would respond to x with a positive primary moral judgment and some with a negative primary moral judgment, then the moral facts about x are *relative*. I suspect the great majority of moral facts are relative, but am unsure whether there might also be a selective elite of objective moral facts, perhaps pertaining exclusively to what is intrinsically, non-instrumentally valuable.

Moral Psychology Again: The Primacy of the Noncognitive

We are now in a position to appreciate the primacy of our noncognitive moral judgments over our cognitive ones. If there were no noncognitive moral judgments, there would be no moral facts, hence no truthmakers for our cognitive moral judgments. The result would be a nasty form of error theory where moral life would consist in so many moral beliefs responsive to *nothing*.

The asymmetry here may be put in terms of the appropriateness or correctness or fittingness of our moral judgments. For our noncognitive moral judgments to be fitting is for them to *conform* to the moral judgments our idealized version would make. For our cognitive moral judgments to be fitting is for them to *represent that* those would be the moral judgments our idealized version would make. Accordingly, there cannot be fitting cognitive moral judgments unless there are noncognitive moral judgments, but there *can* be fitting noncognitive moral judgments even if there are no cognitive moral judgments. It is in this sense that the noncognitive moral judgments are primary and our cognitive moral judgments are secondary.

Secondary moral judgments are still of major importance in our moral life – they are the main source for moral *growth* (at a personal level) and moral *progress* (socially). This is because they present us with the primary moral judgments we *would* have if we were nearer ideal than we in fact are.

Moral Semantics: Descriptivism Despite (Weak) Noncognitivism

We utter moral claims day in and day out. Our public (including political) discourse is in large part a moral discourse. The moral utterances we make, either in private

interpersonal interactions or as participants in the public discourse, could in principle express either our primary or our secondary moral judgments. In practice, the more public our moral utterances, the more they tend to give voice to our *secondary* moral judgments, that is, to our moral *beliefs*. If at a dinner party you utter the sentence 'Eating meat is wrong,' typically you are expressing your belief whose content is <eating meat is wrong>, not your negative moral emotion whose content is <eating meat>. Of course, we do routinely communicate to each other our primary moral judgments. But we tend to do so by *reporting* rather than *expressing* them. At the dinner party, you might say, for instance, 'I am indignant about eating meat,' or 'I feel ashamed when I eat meat,' which report on your mental states. Such utterances do not *give voice* to your mental states, they merely *tell* of them. It is hard to *express* indignation about eating meat – perhaps a combination of 'Yuck!' and the right grimace would do the trick. But this is only partially a *linguistic* expression. Genuine and complete linguistic expression of moral judgment, in such straightforward sentences as 'Eating meat is wrong,' is expression of moral belief.

This is why public moral claims tend to be put in the indicative mood and exhibit a manifest grammar characteristic of descriptive discourse: they *are* descriptive claims, ones whose putative truthmakers are the same as the truthmakers of the moral beliefs they express. Despite my weak noncognitivism about moral judgment, then, I am a descriptivist about moral language.

This descriptivism explains the role of social interaction and public discourse in moral progress. Trafficking in moral beliefs rather than moral emotions, public discourse is a space for processing the question of what moral emotions we *would* have if we were nearer ideal than we are – a question close to that of what moral emotions we *should* have.

Conclusion

The metaethical framework presented above combines weak noncognitivism about moral judgments, weak antirealism about moral reality, and descriptivism about moral discourse. At its heart is a distinction between two kinds of moral judgment: a cognitive

judgment with moral content represented under the guise of the true, and a noncognitive judgment with nonmoral content represented under the guise of the moral good or bad. The noncognitive one is more fundamental, guides moral action in the normal go of things, and is a constituent of moral truthmakers; but the cognitive one is more conspicuous in moral discourse and is more central to moral progress.

References

- Ayer, A.J. 1936. *Language, Truth, and Logic*. London: Victor Gollancz.
- Hare, R.M. 1952. *The Language of Morals*. Oxford: Oxford University Press.
- Horgan, T. and M.C. Timmons 2006. "Cognitivist Expressivism." In T. Horgan and M.C. Timmons (eds.), *Metatethics After Moore*. Oxford: Oxford University Press.
- Kriegel, U. 2012. "Moral Motivation, Moral Phenomenology, and the Alief/Belief Distinction." *Australasian Journal of Philosophy* 90: 469-486.
- Mackie, J.L. 1977. *Ethics: Inventing Right and Wrong*. New York NY: Penguin.
- Smith, M. 1987. "The Humean Theory of Motivation." *Mind* 96: 36-61.
- Stevenson, C.L. 1944. *Ethics and Language*. New Haven: Yale University Press.