

# MORAL MOTIVATION, MORAL PHENOMENOLOGY, AND THE ALIEF/BELIEF DISTINCTION

Uriah Kriegel

In a series of publications, Tamar Gendler has argued for a distinction between belief and what she calls ‘alief’. Gendler’s argument for the distinction is a serviceability argument: the distinction is indispensable for explaining a whole slew of phenomena, typically involving ‘belief-behaviour mismatch’. After embedding Gendler’s distinction in a dual-process model of moral cognition, I argue here that the distinction also suggests a possible (dis)solution of what is perhaps the organizing problem of contemporary moral psychology: the apparent tension between the inherently motivational role of moral judgments and their manifestly objectivistic phenomenology. I argue that moral judgments come in two varieties, moral aliefs and moral beliefs, and it is only the former that are inherently motivating and only the latter that have an objectivistic phenomenology. This serves to both bolster the case for the alief/belief distinction and shed new light on otherwise well-trodden territory in metaethics. I start with an exposition of the moral-psychological problem (§1) and a discussion of Gendler’s alief/belief distinction (§2). I then apply the latter to moral judgments in an attempt to dissolve the former (§3). I close with discussion of the upshot for our understanding of moral thought, moral motivation, and moral phenomenology (§4).

**Keywords:** cognitivism, internalism, alief, besire, dual-process

## 1. The Moral-Psychological Problem

In the opening chapter of his book *The Moral Problem*, Michael Smith [1994] articulates the organizing problem of contemporary metaethics, as he sees it, in terms of a triad of theses which are all highly plausible and yet hard to reconcile. Smith’s own triad is not formally inconsistent, but is readily adapted into one that is:

- 1) Cognitivism: Moral judgments purport to be about objective matters of fact.
- 2) Internalism: Moral judgments are inherently motivational.
- 3) Humean Psychology: Mental states, including moral judgments, cannot both have objective purport and be inherently motivational.<sup>1</sup>

<sup>1</sup>The metaethical literature often uses the term ‘moral judgment’ to denote not a certain type of mental state, but a certain type of utterance. (Indeed, this is the case with Smith’s own discussion.) Here I use the term to denote a type of mental state.

Smith calls his triad ‘the moral problem’. As this triad is slightly different, I will call it ‘the moral-psychological problem’.<sup>2</sup>

Smith does not quite offer a positive argument for the cognitivist thesis in his triad, but only defends it against counter-arguments.<sup>3</sup> Presumably, this is because he takes cognitivism to be antecedently plausible. It is an interesting question, however, what *makes* it antecedently plausible. One answer is that it is simply *intuitive* that moral judgments have objective purport. Another answer, which seems to go deeper, is that the objective purport of moral judgments is *phenomenologically* manifest, in that what it is like for a subject to make a moral judgment often involves a feeling as of homing in on an objective matter of fact.<sup>4</sup> The advantage of appealing to phenomenology rather than intuition here is that an objectivistic phenomenology would explain *why* we have the intuition, whereas the intuition would not explain any phenomenology. Thus appeal to moral phenomenology as the grounds for moral judgments’ objective purport involves added explanatory depth.

To be sure, positive arguments for cognitivism, perhaps of a more theoretical nature, may be available as well. Probably the best known is an argument from the Frege–Geach problem. The problem is that if moral judgments were *not* cognitive states, they could not play the kind of inferential role in reasoning and deliberation—which includes inferential interaction with cognitive states—that they appear to [Geach 1960].<sup>5</sup> Various more or less technical responses to this argument are available (see Gibbard [2003] for a particularly ingenious response), but in any case the deep source of plausibility of cognitivism does not seem to be such theoretically sophisticated considerations. Cognitivism just *seems* right, and the source of this seeming, I contend, is something like the phenomenology of moral judgments.

Smith [1994: Ch.3] does offer a positive argument for the internalist thesis in his triad. The argument has been widely discussed, and I will only sketch

<sup>2</sup>Smith’s own triad of theses is this: 1) Moral judgments express beliefs about objective matters of fact (the ‘objective pretensions’ of moral judgment); 2) If a person makes a moral judgment, then *ceteris paribus* s/he is motivated to act accordingly (the ‘practicality’ of moral judgment); 3) Action can be caused only by a combination of a means-end belief and a desire for the relevant end (‘Humean psychology’).

<sup>3</sup>In particular, he focuses on a destructive dilemma adapted from Ayer [Smith 1994: Ch. 2]. The dilemma concerns whether the putative objective facts allegedly targeted by moral judgments would be natural or non-natural facts. If non-natural, then the *a priori* supervenience of moral facts on natural facts would be utterly inexplicable. If natural, a second dilemma arises, concerning whether moral facts’ being natural is an *a priori* or a *posteriori* truth. On the one hand, this cannot be *a priori*, because there are no complete analyses of moral terms in natural terms; but it cannot be *a posteriori* either, because that would make it impossible to recover certain platitudes about morality. Smith chooses the natural horn of the dilemma, claiming that, *a priori*, the facts allegedly targeted by moral judgments are natural facts; this apriority does not require complete analyses of moral facts in natural terms, he insists, but only what he calls ‘summary definitions’.

<sup>4</sup>Some may be squeamish about talk of phenomenology in the context of discussion of moral judgments, perhaps judgments in general. I have no such hang-ups: the recent literature in the philosophy of mind is replete with arguments for *cognitive* phenomenology—a proprietary phenomenology of belief and judgment [Moore 1953; Cohen 1992; Strawson 1994; Peacocke 1998; Siewert 1998; Horgan and Tienson 2002; Pitt 2004; Kriegel 2003, forthcoming]—and even sustained arguments and detailed discussions of a specifically *moral* phenomenology, a phenomenology proper to our moral mental life [Scheler 1913; Mandelbaum 1955; Horgan and Timmons 2005, 2007, 2008; Kriegel 2008, 2008a, forthcoming].

<sup>5</sup>For example, one can reason as follows: if my sister did not visit my father in hospital yesterday, then I ought to do so today; she did not; therefore, I ought to. This type of inferential interaction appears to impute a cognitive/descriptive content. If we construe moral judgments as having a conative/imperative content, they would not integrate as well into this form of reasoning, since the following seems ill-formed: if my sister did not visit my father in hospital yesterday, then visit him today!; she did not; therefore, visit him today!

its structure in the barest outlines here. Smith observes that, as a matter of contingent fact, judgments that one ought to  $\varphi$  and motivations to  $\varphi$  are reliably compresent in the good and strong-willed person. This fact requires explanation. The most straightforward explanation is that there is an ‘internal’, constitutive connection between moral judgment and motivation—that moral judgments are ‘practical’, in the sense that they are inherently motivating.<sup>6</sup> The alternative explanation would have to be that there is a merely ‘external’, causal connection between the two, owing to the fact that the good and strong-willed person tends to have a psychologically real desire to do whatever it turns out she ought to do. This alternative explanation is implausible, according to Smith, because it is implausible to ascribe such a desire to all good and strong-willed persons.<sup>7</sup> This is effectively an argument from inference to the best explanation for internalism: the *explanandum* is the reliable compresence of moral judgment and corresponding moral motivation in the right kind of agent, and the best *explanans* is that moral judgments are inherently motivating, that is, that there is a constitutive connection between moral judgment and moral motivation.

Unfortunately, the objective purport of moral judgments and their inherently motivational nature are uncomfortable bedfellows. The tension between them can be brought out through what Smith calls ‘Humean Psychology’. This is the traditional idea that the mind divides into two independent departments (independent metaphysically, not causally)—reason and the will, the cognitive and the conative, or what have you—and mental states belong to either one or the other but not both. The problem is that objective purport characterizes one type of mental state, the cognitive one, whereas motivational role characterizes the other, conative type. The core of Humean Psychology is the claim that there are no unitary mental states which boast both inherent motivational force and objective purport, states both cognitive and conative.<sup>8</sup> Such states are sometimes referred to as ‘besires’—half belief, half desire—and Smith argues that the notion of besire is incoherent. It is therefore incoherent to be a cognitivist internalist—to hold that moral judgments are besires.<sup>9</sup>

What is the rationale for Humean Psychology and its ban on besires? Commentators have focused on Smith’s argument from ‘direction of fit’, but this focus strikes me as misguided. The direction-of-fit argument is that mental states cannot involve both the belief attitude and the desire attitude towards the same content, because it is a conceptual truth that when the state of affairs making up the state’s content goes out of existence, the desire attitude tends to persist whereas the belief attitude tends to desist. The

<sup>6</sup>It should be remembered that the term ‘internalism’, as originally used by Falk [1945] and Frankena [1958], was chosen precisely to intimate such an ‘internal’ connection.

<sup>7</sup>This is not because it is more plausible to ascribe to them the opposite desire, but because it is more plausible to ascribe to them no relevant desire.

<sup>8</sup>This raises the question of what is meant by a mental state’s being ‘unitary’, something that is not elucidated all that satisfactorily in Smith’s discussion but will be partially addressed in this article.

<sup>9</sup>Prominent versions of such cognitivist internalism are developed by McDowell [1979] and McNaughton [1988].

problem with this argument is that philosophers sympathetic to besires typically do not hold that moral judgments involve the belief attitude and the desire attitude *towards the same content*. Rather, they typically hold that a moral judgment that one ought to  $\varphi$  involves a belief attitude towards the content ‘I ought to  $\varphi$ ’ and a desire attitude towards the content ‘I  $\varphi$ ’. Thus their view is unthreatened by the argument from direction of fit. What threatens it is rather Smith’s (subsidiary) argument from *modal separability*: the belief that one ought to  $\varphi$  and the desire to  $\varphi$  are modally separable in that each could exist in the other’s absence, which means that they do not make up a unitary mental state. And it is precisely the point of Humean Psychology that beliefs and desires are distinct states rather than unitary ones.

Two clarifications are in order regarding Humean Psychology and its ban on besires. First, the ban is not on any propositional attitude with a cognitive component and a conative component that can be denoted with a simple term. For example, being disappointed that  $p$  involves believing that  $p$  and desiring that  $q$  (where  $q = \sim p$ ), but disappointment is not a besire. What distinguishes disappointment from besire is precisely the modal separability of its components, which casts it as a mere sum of two elements rather than an organic whole.<sup>10</sup> Essentially, a disappointment involves not only two different contents, but at bottom also two different attitudes, whereas a besire involves a single, non-decomposable attitude.<sup>11</sup> Secondly, drawing an exclusive (though not exhaustive) distinction between cognitive and conative states does not mean casting cognitive states as motivationally impotent. To say that belief is not ‘inherently motivating’ is not to say that it is *epiphenomenal*. Rather, it is to say that there is only an ‘external’, contingent connection between belief and motivation, as opposed to the ‘internal’, necessary connection desire exhibits (and other conative states). That the connection of belief to motivation lies in between the two extremes of epiphenomenalism and internal/necessary connection strikes me as precisely the *prima facie* plausible position on the motivational character of belief.

Thus all three theses comprising the moral-psychological problem have much to recommend them. Yet they cannot all be true. Arguably, this is indeed the organizing problem of contemporary moral psychology, at least since Mackie’s [1977] groundbreaking *Ethics: Inventing Right and Wrong*, which explicitly pitted these theses, or ones very like them, against each other. The problem can be solved straightforwardly if we reject Humean Psychology. But many have been impressed with the arguments for it, thus conceding a genuine tension between objective purport and motivational

<sup>10</sup>The distinction between ‘mere sum’ and ‘organic whole’ may be cast in terms of Simons’s [1987] mereological distinction between sums and complexes. Intuitively, the idea is that complexes involve an essential connection between the elements, whereas a sum is but the mere compresence of those elements; a pre-assembled IKEA desk is a sum, an assembled one a complex. More formally, a whole  $W$  of parts  $P_1, \dots, P_n$  is a sum just in case the only way to destroy  $W$  is to destroy a  $P_i$ ;  $W$  is a complex just in case it is possible to destroy  $W$  without destroying any  $P_i$  (namely, by destroying the essential connection among  $P_1, \dots, P_n$ ).

<sup>11</sup>It is worth noting that the decomposition of attitude is not an artefact of the presence of two different contents. For it can occur even when the mental state involves a relation to a single content. Thus, being glad that  $p$  involves believing that  $p$  and desiring that  $p$  (or at least some cognitive attitude towards  $p$  and some conative attitude towards  $p$ )—but gladness is not a besire either.

nature, cognitivism and internalism, moral phenomenology and moral motivation. Contending with all the dialectical pressures in the area has recently led to a number of positions offering rather convoluted portraits of moral judgment. Copp's [2001] 'realist expressivism', Kalderon's [2005] 'hermeneutic moral fictionalism', and Horgan and Timmons's [2006] 'cognitivist expressivism' are three prominent examples; there are others. The merits and demerits of such views deserve separate discussion. Here I propose that a neat, unconvoluted portrait of moral judgments is offered by recent developments in the philosophy of action, backed by advances in cognitive-scientific models of cognitive architecture. The next section introduces these developments.

## 2. Dual-Process Theory and the Alief/Belief Distinction

There is a long philosophical tradition of theorizing the operation of the faculty of thought in the image of logic: thought is taken to be guided by deductive and inductive logic, such that, at least in the well-functioning thinker, causal connections between mental states somehow follow logical connections (deductive or inductive) between these states' representational contents. There is also, however, a more recessive strand in philosophical theorizing about thought, dating back at least to Hume, that offers an alternative model: far from following logical relations, causal connections among thoughts are seen as determined by certain *laws of association* among their contents. Hume himself offered three such laws—resemblance, contiguity, and causality—but his followers among eighteenth- and nineteenth-century British associationists, from David Hartley to G. H. Lewes, refined and developed these laws more systematically. Thereafter emerged a debate between 'rationalist psychology' and 'associationist psychology' as to which model captures better the actual operation of the faculty of thought.<sup>12</sup>

So which is it: does thought operate along rationalist lines or associationist ones? Here is an interesting answer: both! There are really two complementary faculties operative in us, one conforming to the rationalist model and one to the associationist model. A classical statement of this two-system approach is offered by Sloman [1996: 3]:

One system is associative because its computations reflect similarity structure and relations of temporal contiguity. The other is 'rule based' because it operates on symbolic structures that have logical content and variables and because its computations have the properties that are normally assigned to rules. These systems serve complementary functions and can simultaneously generate different solutions to a reasoning problem.

As it turns out, this dual-process architecture is quite common in the cognitive system: dual-process models have been offered, and have

<sup>12</sup>This served as an obvious precursor to the modern-day debate between proponents of 'classical architecture' and of 'connectionist architecture'.

gained much traction, for memory [Barrett *et al.* 2004], learning [Sun *et al.* 2005], and decision-making [Kahneman 2003].<sup>13</sup> If this is such a general architectural feature of the human mind, it stands to reason that it might characterize the faculty of thought as well. And indeed Sloman and others have offered wide-ranging empirical evidence suggesting it does.

From a design standpoint, the rationale for a dual-process architecture is likely to do with the inescapable speed-for-accuracy trade-off in cognition.<sup>14</sup> In many cases, the cognitive system is well served by employing two distinct mechanisms that hit different trade-off optima. In the normal go of things, and as long as things go fine, the system defaults to the fast-but-inaccurate ('quick and dirty') mechanism; it is only when a problem or conflict arises, or the stakes are particularly high, that the system is liable to shift to the accurate-but-slow mechanism. Note that if this is the teleological rationale for dual-process architecture, it again stands to reason that thought would involve such an architecture, using by default an associationist mechanism good enough for the purposes of everyday life and reverting to the rationalist, rule-based mechanism only when greater accuracy is required.

Wherever there is a duality of *processes*, there is likely also a duality of *products*. Thus the associationist system likely produces one mental state, the rationalist system another. It is probably the wrong question to ask whether these two products are of the same 'general kind'. Kind generality comes in degrees: a dog and a hamster belong to the same general kind at one level (they are both mammals) but not at another (they are not both rodents). Likewise, the products of the associationist and rationalist systems are probably of the same kind at one level of generality but not at another. The more appropriate question is therefore this: just *how* similar and dissimilar are the products of the two systems?

In the psychological literature, the two systems are typically described as producing competing *judgments* concerning the experimental task. However, not much conceptual illumination is offered in terms of defining what is involved in a mental state being a judgment, and no systematic conceptual distinction has been drawn between the different *kinds* of judgment produced. For this we are more likely to be helped by philosophers. And indeed, in the work of Tamar Szabó Gendler [2008a, 2008b] we can find a conceptual distinction between two kinds of mental state, plausibly described as judgments, that I contend can play the theoretical roles of 'product of the associationist system' and 'product of the rationalist system'. To my knowledge, Gendler herself does not cast the two states in this way, and I am not sure how sympathetic she would be

<sup>13</sup>There are also dual-process accounts of 'simpler' mental phenomena, including vision [Milner and Goodale 1995]. But it is not entirely clear that the notion of 'dual-process' is exactly the same in those cases.

<sup>14</sup>The cognitive system always faces the question of how much of its finite time and energy resources it should assign to the task facing it; the fewer resources it decides to expend, the less competently it will perform the task. In many cases, this will translate into less accurate representation of the environment. This is, moreover, perfectly 'rational': to serve the organism's overarching purposes (reproductive success, presumably), cognition in the wild is better off as a satisficer than as a maximizer of accuracy. Accuracy is the first casualty of war of survival.

to placing her distinction within the context of dual-process models. Nonetheless, I contend that the two are natural allies and can be combined to articulate a picture of thought and reasoning that is compelling both empirically and conceptually.

Gendler distinguishes between belief and what she calls *alief*. The distinction can be appreciated either through examples or through a more theoretically informed elucidation of the notion of alief. Gendler herself offers myriad examples; here is a particularly simple one: unless one is particularly reflective regarding such matters, when one sees a sign exclaiming ‘glasses: \$9.99’, one tends to form the belief that the glasses cost \$10 but the alief that they cost \$9.<sup>15</sup> When one excitedly hurries to buy the glasses, even though one is in need of none, in all likelihood it is one’s alief that guides one’s action.

Consider another example. Every day I bike to my office. Today, however, they are predicting rain. (This virtually never happens around here.) Hearing this, I realize that I should take the car. I eat my breakfast, brush my teeth, reply to a couple of emails, take my notes, and head out. Once out the door, I assertively head towards the bike shed in my yard, never for a moment stopping to think that the car is parked on the street. It is not plausible to say that I believe the car is in the shed, nor that I believe I should bike to campus. No: I believe I should take the car, and believe it is not in the shed. My belief mismatches my behaviour—which frustrates causal explanation of the latter in terms of the former. What causally explains my behaviour is that I *alieve* I should bike to campus. This alief has been formed through years of pleasant routine associating mornings, sheds, bikes, campus, etc. Arguably it causally explains my shed-bike-retrieval behaviour *most* mornings, even on days my alief and belief converge, but is appreciated to do so only thanks to the occasional belief-behaviour mismatch.<sup>16</sup>

As for the more theoretical characterization of alief, this can be offered either through specification of alief’s *explanatory role* in the theory of action or through a more substantive characterization of the likely *occupant* of said role. The *role* appears to be this: an alief is a mental state whose occurrence causally explains our behaviour in cases where our behaviour does not match our beliefs, but may be operative as well when our behaviour does match our beliefs. What exactly *occupies* this explanatory role is something Gendler is somewhat tentative about and offers only a ‘provisional’ characterization of: she proposes that an alief is a mental state whose content consists in a cluster of representational, affective, and motivational components ‘associatively linked’ (i.e., tending to co-occur). Thus, the full content of my alief about the glasses might be (expressed by) something like ‘\$9 for glasses. Exciting! Get them!’ This has three components: the first

<sup>15</sup>More precisely, as we will see in a moment, the content of the alief features also affective and motivational components.

<sup>16</sup>I take this to indicate that alief is normally the genuinely causal-explanatory state. However, it should be noted that the inference from claiming that alief is the operative state in case of alief-belief disharmony to claiming that alief is operative even in cases of alief-belief harmony is non-trivial and would ultimately require empirical substantiation. At present I take this to be a plausible speculation.

(‘\$9 for glasses’) is representational, or ‘cognitive’; the second (‘Exciting!’) is affective; the third (‘Get them!’) is motivational.<sup>17</sup>

Gendler’s case for the alief/belief distinction is a sort of argument from serviceability: the distinction is indispensable for explaining a host of mental phenomena that would be utterly perplexing without it. The glass-purchase and bike-retrieval behaviours are cases in point, but Gendler offers a wide and varied range of cases that call for a unified explanation in terms of causally operative aliefs.<sup>18</sup> Thus the hypothesis that there are two kinds of mental state plausibly described as ‘judgment’—alief and belief—enjoys explanatory scope and depth that render it highly credible.

To this consideration of explanatory scope and depth I would add one more layer of theoretical virtue: explanatory *unity*. Given the prevalence of dual-process models in cognitive science, another model positing a duality of causally operative mechanisms in a single cognitive domain, a belief-producing mechanism and an alief-producing mechanism, integrates smoothly into a satisfyingly unified account of human cognition and action. Indeed, if we hypothesize that the alief-producing mechanism is one and the same as Sloman’s associationist system and the belief-producing mechanism one and the same as his rationalist (‘rule-based’) system, the explanatory unification and integration of empirical evidence and firm conceptual underpinnings becomes extraordinarily powerful.

Understanding the alief/belief distinction in the context of dual-process/two-system models in cognitive science also helps us claim certain important differences between the two more confidently. In such models, the two processes or systems are typically divided into a Low Road and a High Road. The low-road system is fast, efficient, automatic, implicit, unconscious, specialized, and directly tied to action, but not very flexible, not particularly amenable to monitoring and control (because informationally encapsulated), and with limited range. The high-road system is the converse: flexible, wide-ranged, consciously and explicitly controllable and manipulable, but linked to action only indirectly, slow and inefficient. We can expect the same differences to apply to the alief- and belief-producing mechanisms. An *alief* would be produced by processes that are typically automatized, relatively uncontrollable, architecturally connected to motor output, and inaccessible to consciousness. A *belief* would be produced by more deliberate and controllable processes that have a more indirect impact on action but are often conscious.<sup>19</sup>

Much more can be said about the notion of alief and the distinction between alief and belief as two different kinds of judgment—and doubtless much more *will* be said, by philosophers and cognitive scientists alike, in coming years. My hope is that the present discussion has given the notion of alief sufficient texture to proceed to the main mandate of this paper: the application of the alief/belief distinction to *moral* judgment.

<sup>17</sup>In the bike-retrieval scenario, perhaps my alief’s full content is ‘Bike in the shed. How reassuring! Retrieve and use!’

<sup>18</sup>I do not rehearse these here—see Gendler’s [2008a, 2008b] work for discussion.

<sup>19</sup>Recall from §1 that belief’s indirect connection to action does not amount to causal inertia, but only to a merely contingent or ‘external’ causal connection with action.

### 3. Moral Judgments, Moral Beliefs, and Moral Aliefs

The need for a distinction between belief and alief does not depend on any particular subject matter: judgments about anything below or above the moon may have the characteristics of belief or of alief. Thus we have every reason to expect that moral judgments can also come in two varieties: moral aliefs and moral beliefs. This expectation is confirmed by both types of consideration used to support the alief/belief distinction in general: (i) its explanatory indispensability *vis-à-vis* cases of belief-behaviour mismatch and (ii) its explanatory unity with ‘the best science of the mind we got’ (as Fodor used to say).

On the one hand, Gendler-style examples of belief-behaviour mismatch are easy to come by in the moral domain. Discrepancy between a person’s honest proclamations of moral principle and their ongoing moral practice is a familiar phenomenon. As ‘implicit bias’ studies have consistently shown, many of us who would never consciously assent to a racist proposition, such as ‘Black men are generally more dangerous than white men’, nonetheless show stubborn traces of racist dispositions in our behaviour, a tendency registered in differential activation of the amygdala (associated with threat detection) in the presence of startling white and black faces [Amodio *et al.* 2003].<sup>20</sup> It is tempting to describe this precisely as a situation in which we have no racist beliefs but some racist aliefs, as Gendler [2008b] indeed does.

This kind of ‘negative’ belief-behaviour mismatch—where a blameworthy moral attitude lies hidden beneath the surface of a praiseworthy attitude—is routinely pointed out in the public and private spheres. Its converse, ‘positive’ mismatch receives much less attention but is a common occurrence as well. Arpaly [2002] writes of her undergraduate student who brandishes her copy of Ayn Rand’s *Atlas Shrugged* and passionately extols the virtues of selfishness, all the while helping her peers to prepare for exams, volunteering at the local soup kitchen, and so on. It is tempting and plausible to describe this student as holding egoistic beliefs but altruistic aliefs.

As for explanatory continuity with cognitive-scientific research, given the prevalence of dual-process architecture in our cognitive system, it is natural to hypothesize that such a model might apply to moral cognition as well. The idea is that the execution of moral-cognition tasks is subserved by two independent neurophysiological and/or computational underlying mechanisms, an associationist low-road mechanism and a rationalist high-road mechanism. Presumably, these two mechanisms would produce mental states that would bear characteristic properties: the former would produce moral aliefs, the latter moral beliefs.

As a matter of fact, dual-process models of moral cognition have flourished in the past decade, as several cognitive and social psychologists have been more or less independently led to posit them (see Greene and Haidt [2002] for a classic articulation, with discussion of pertinent experimental results). These dual-process models are typically *not* billed as distinguishing a

<sup>20</sup>Thus the phenomenon is supported by both behavioural and neural evidence—the two kinds of evidence typically appealed to in cognitive science. Needless to say, it is also familiar from everyday life.

430 rationalist system from an associationist one, but as distinguishing a  
 435 rationalist system from a *sentimentalist* one. Thus whereas the high road is  
 taken to be rational and rule-based, the low road is cast as sentimentally  
 grounded, driven by (often socially conditioned) affective responses  
 evolutionarily preceding any capacity for rational processing. However, it  
 440 is an open question whether these models could be reconceptualized in  
 associationist-cum-alief-ish terms without loss of predictive or explanatory  
 power (and potentially with some gain therein). For one thing, the low-road  
 sentimentalist process in these models is typically theorized as habitual, fast,  
 efficient, etc., just as the alief-forming associationist mechanism is supposed  
 to be. Furthermore, like aliefs, the mental states the sentimentalist typically  
 445 appeals to have primarily affective content but feature representational and  
 motivational elements as well. This suggests to me that sentimentalist  
 construals of low-road moral cognition are likely amenable to subsumption  
 under an associationist construal, with the associationist system cast as  
 trafficking heavily in ‘sentimental states.’<sup>21</sup>

450 This is not the place to launch such a project of subsumption. The  
 preceding discussion—in particular, of the cases of implicit racism and  
 Arpaly’s student—should suffice to create a presumption in favour of the  
 applicability of the alief/belief distinction to moral judgments, hence in  
 favour of a distinction between moral aliefs and moral beliefs. These exhibit,  
 one would think, the characteristic features of aliefs and beliefs in general. A  
 moral alief is produced by uncontrollable and automatized processes that are  
 architecturally connected to motor output and inaccessible to consciousness.  
 A moral *belief* is produced by more deliberate and controllable processes that  
 455 have a more indirect impact on action but are often conscious.

460 It is worth pausing to dwell on the differential connection of moral alief and  
 belief to action and to consciousness, as displayed in the normal unreflective  
 subject. As operative in everyday life, moral aliefs are formed by sub-personal  
 processes that are never conscious. This may sometimes result in aliefs that  
 are themselves unconscious, but often the alief itself is conscious. (Compare:  
 465 visual processes are typically unconscious, but their products—visual  
 experiences—are typically conscious.) However, much more essential to  
 aliefs is their action-guiding profile. As we have seen, aliefs govern behaviour  
 as it freely unfolds in the normal go of things, where explicit deliberation is  
 not called for. Thus, Arpaly’s supposedly Randian student may proclaim  
 disdain for altruism, but what guides her action are by and large her altruistic  
 470 aliefs, which in the normal go of things trump her egoistic beliefs. This direct  
 connection to behaviour is, moreover, *constitutive* of alief: a mental state  
 would not *be* an alief if it were not architecturally connected to motor output  
 in a way that enables direct impact on ongoing behaviour; it is part of what  
 makes an alief an alief that it is connected to behaviour in the requisite way. In  
 this respect, alief is an inherently motivational state *par excellence*.

Moral beliefs, by contrast, are dissociated from ongoing action; their  
 impact on action is much more indirect, often concerning the long term.

475 <sup>21</sup>Short of subsumption, the case for a distinction between a sentimentalist low road and a rationalist high  
 road could inspire an *analogous* case for a distinction between an associationist low road and a rationalist  
 high road.

Perhaps after years in the cradle of Randian ideology, undergoing Aristotelian re-education of her behavioural habits, Arpaly's student could slowly become a better egoist. But this would only be because her beliefs would have finally refashioned her aliefs in their own image. It is relatively rare that a belief trumps an alief in the causation of behaviour.<sup>22</sup> It may happen on occasion that, in conjunction with some desire, a moral belief may guide action over the objections, so to speak, of a contradictory moral alief (e.g., when the Randian student resists helping a peer on some particular occasion on ideological grounds). But among the innumerable actions we perform in the course of a day, many more are caused by aliefs.<sup>23</sup> And in any case, as the products of a high-road system, moral beliefs are not *architecturally* connected to action in the way moral aliefs are. What characterizes moral beliefs more centrally is their conscious accessibility: they 'live' in personal-level cognition, where the subject typically has immediate, non-inferential, first-person awareness of them and their contents.<sup>24</sup> As a result, the phenomenal profile of beliefs—the subjective feel of being aware of a proposition and presenting it to oneself as true or plausible—is much more essential to them. This presenting of something to oneself as true or plausible is characteristic of moral beliefs as much as other beliefs: it *feels true* that involuntary servitude is wrong.<sup>25</sup> From the inside, believing that involuntary servitude is wrong, as opposed to merely entertaining it, feels just as objectively compelling as believing that water is H<sub>2</sub>O (as opposed to merely entertaining it). This is the objectivistic phenomenology of moral belief mentioned in §1.<sup>26</sup>

<sup>22</sup>By this I mean that it is much more common for the alief to trump the belief than for the belief to trump the alief. (I do *not* mean that the trumping of alief by belief is rare by some independent standard.) That this is so falls out of the dual-process picture presented here, but is also independently plausible: as we consider ongoing behaviour in unreflective mode throughout the day, we observe that aliefs normally 'run the show', beliefs supplanting them only when some conflict arises in the system, the sort of conflict that engages so-called cognitive control mechanisms (see empirical research on the Stroop Effect).

<sup>23</sup>In saying this, I do not mean to intimate that the matter of individuating and counting actions is trivial. But however issues about action counting are resolved, it will turn out that many more are caused by aliefs than by beliefs.

<sup>24</sup>Such first-person awareness will characterize only *conscious* moral beliefs, and not all moral beliefs are conscious, certainly not at any time. But it is in the essence of belief (as seen in the present picture) to be *accessible* to consciousness, that is, be disposed to readily become conscious. So, more precisely, we may say that it is in the essence of moral beliefs to be disposed to be readily lend themselves to first-person awareness.

<sup>25</sup>Cohen [1992: 4, 11] articulates the central idea rather nicely (despite being unrecognized in the cognitive phenomenology circles):

... belief that *p* is a disposition ... normally to feel it true that *p* and false that not-*p*, whether or not one is willing to act, speak, or reason accordingly ... Feeling it true that *p* may thus be compared with feeling it good that *p*. All credal feelings, whether weak or strong, share the distinctive feature of constituting some kind of orientation on the "True or false?" issue in relation to their propositional objects, whereas affective mental feelings, like those of anger or desire, constitute some kind of orientation on the "Good or bad?" issue.

<sup>26</sup>The case can be made in much more detail by adapting a phenomenological analysis of belief due to Horgan and Timmons [2007], who proceed by isolating five phenomenological characteristics of belief in general and showing that some moral judgments exhibit those characteristics as well. Beliefs (a) involve a feeling as of 'coming down' on an issue, in which (b) there is an application of a sortal, or a categorization of objects, in a manner that feels (c) involuntary, (d) rationally imposed by consideration of reasons for believing, and (e) lending itself to expression via a declarative sentence. These characteristics are not necessarily very vivid or obvious—they may well reside in the phenomenal periphery or 'fringe of consciousness'—but they are phenomenologically real nonetheless, and are present in such moral judgments as are produced by the high-road process of moral cognition.

525 If all this is right, then the distinction between moral aliefs and beliefs may hold the potential for a dissolution of the moral-psychological problem. The theses in the inconsistent triad with which we started concern ‘moral judgment’, but in the light of this section’s discussion, that term is seen to be potentially equivocal: in one sense, it might denote moral belief; in another, moral alief. (I say ‘potentially equivocal’ because, strictly speaking, the term, if used along the lines of terminological regime proposed here, would be a *univocal* term denoting the genus of which moral belief and alief are both species. In practice, however, the recommended terminological regime need not be in effect, and the term could be used equivocally to denote either one species or the other. As it happens, I think this is exactly what is going on.)  
 530  
 535 Once we disambiguate ‘moral judgment’, I will now argue, we find that there is no single sense of the term in which all three theses are compelling.

Recall that the cognitivist thesis in the triad is a thesis about the phenomenology of moral judgment: that it has an objective purport. As we have just seen, this is highly plausible for moral *beliefs*. But is it plausible for moral aliefs? Consider a racist moral alief of the sort we have been discussing. Its content may be thought of as ‘Black man. Scary! Must get away!’ There is in fact a representational or cognitive component in this content (‘Black man’), and this component does involve a phenomenology as of being directed at an objective matter of fact. But this component has no moral or normative dimension to it at all—it is purely descriptive. It is the affective and motivational components (‘Scary! Must get away!’) that involve a normative or moral dimension.<sup>27</sup> What this means is that although a moral alief has a phenomenology as of being directed at an objective matter of fact, it does not have this phenomenology *qua moral*. What makes it moral and what makes it have this phenomenology are separate components, indeed components that are not unified in the way prohibited by Humean psychology. In this respect, an alief is more like disappointment than like desire: the alief involves a cognitive attitude towards one content and a conative attitude towards another, where the attitudes are modally separable.  
 540  
 545  
 550  
 555

Recall, next, that the internalist thesis in the triad requires moral judgments to be inherently motivating. Once we draw the alief/belief distinction, however, we see that this is antecedently plausible only for moral *aliefs*. Moral beliefs, once clearly distinguished and divorced from moral aliefs, are freed from this theoretical burden, and seen to be more removed from action and motivation. Although on occasion they may conspire with some pro-attitude to bring about an action, and have a role in the long-term shaping of aliefs, they are not *architecturally* linked to action and thus are not *constitutively* motivational states.  
 560

565 It would seem, then, that we can comfortably and in good conscience adopt only two of the three theses comprising the moral-psychological problem for each disambiguation of ‘moral judgment’: cognitivism and Humean Psychology about moral beliefs, internalism and Humean Psychology about moral aliefs. In other words, we can reject internalism  
 570

<sup>27</sup>Certainly many moral philosophers would want to say that the concept of the scary, or dangerous, is a thick moral concept; and that ‘Must get away!’, with its imperative content, has the hallmarks of action-guiding normativity.

about moral beliefs (they are not constitutively, or architecturally, connected to action) and we can reject cognitivism about moral aliefs (they do not, *qua moral*, purport to be about objective matters of fact). Thus there is no single notion for which all three theses in the triad are compelling. 575

The triad can now be dissolved into two perfectly valid (and plausibly sound) arguments about the nature of different kinds of moral judgment. The first is an argument about the nature of moral aliefs: 580

- 1) Moral aliefs are inherently motivating.
- 2) Mental states in general, including moral aliefs, cannot both have objective purport and be inherently motivating. 585

Therefore,

- 3) Moral aliefs do not purport to be about objective matters of fact.

The second argument concerns moral beliefs: 590

- 1) Moral beliefs purport to be about objective matters of fact.
- 2) Mental states in general, including moral beliefs, cannot both have objective purport and be inherently motivating. 595

Therefore,

- 3) Moral beliefs are not inherently motivating.

The emerging picture combines cognitivism and externalism about moral beliefs with internalism and non-cognitivism about moral aliefs—while holding on to Humean Psychology. My contention is that this picture does justice to all the pre-theoretical and dialectical pressures in the area, and integrates smoothly into an independently plausible (both empirically and conceptually) account of judgment in general. Its upshot is that both cognitivism and internalism are true of moral judgment, but in different senses. We may call this picture *dual-process cognitivist internalism*. 600  
605

#### 4. Moral Motivation and Moral Phenomenology 610

The picture avoids the Frege–Geach problem quite elegantly. Moral beliefs play the kind of inferential role Geach demanded from moral judgments in reasoning and deliberation. Moral aliefs do not, but nor should they be expected to: the cognitive processes they are implicated in are not serial, rule-based bits of reasoning, but associationist, parallel ones. In any case, the fact that *some* moral judgments—the belief ones—are cognitive explains why these sorts of reasoning can and often do take place (this is the Frege–Geach ‘datum’). 615  
620

Meanwhile, the picture affords a more precise take on Smith's argument for internalism about moral judgments—the argument from best explanation of the reliable compresence of judgment and motivation in the good and strong-willed person. As we can now appreciate, 'judgment' is potentially ambiguous, so two different *explananda* may be at play here. First, there is the reliable compresence of motivation and *alief* in the good and strong-willed person. This is best explained by the fact that moral aliefs are constitutively connected to motivation, which indeed they are on the picture at hand. Secondly, there is the reliable compresence of motivation and *belief* in the good and strong-willed person. This is best explained by the fact that the good and strong-willed person would presumably bring her aliefs and beliefs into alignment. The implicit racist is not as good as she could be, and Arpaly's student not as strong-willed. In the good and strong-willed person, these sorts of divergence between alief and belief evaporate.

A common challenge to internalism is the argument from the conceivability of the amoralist. It is conceivable that an agent should realize that she ought to  $\varphi$  and yet fail to come up with the motivation to actually  $\varphi$ —perhaps she is depressed, perhaps she just doesn't care enough [Brink 1989]. There are various more or less technical responses to this challenge—most notably, 'inverted comma' responses [Hare 1952]—but the present picture neutralizes the challenge rather straightforwardly. In conceiving of the amoralist, we conceive of someone who *believes* that they ought to  $\varphi$  but who fails to *alieve* that they ought to  $\varphi$ : someone who has reasoned her way to the realization that  $\varphi$ -ing would be the right thing to do, but whose motivational set-up, determined by associative habit, has not yet caught up with the realization. Thus in conceiving of the amoralist we conceive of someone whose moral beliefs are disconnected from motivation, in a way that suggests that moral beliefs are not *constitutively* connected to motivation; but none of this undermines the notion that moral *aliefs* are constitutively so connected. Thus dual-process cognitivist internalism accommodates the conceivability datum straightforwardly.

Perhaps the only dialectical pressure in this general area that still has any force as applied to dual-process cognitivist internalism is the anxiety surrounding 'error theory'. Moral beliefs, in the present picture, purport to be about objective matters of fact. If, as is not implausible, there simply are no objective matters of fact of a moral nature for beliefs to be answerable to—if there are no moral truthmakers, in short—then moral beliefs would have to be declared erroneous one and all. Some are willing to bite the bullet [Mackie 1977], others are happy to posit moral facts and truthmakers (moral realists). But many, including me, are uncomfortable with either of these moves.

It is important to appreciate, however, that error theory is significantly less disconcerting in the present picture than in standard cognitivism. For in the present picture error theory about moral belief would not undermine the rational foundations of our entire moral practice, since the formation of moral beliefs is only one part—and not the major part—of the practice. In Mackie's picture, *all* our moral judgments are beliefs, and therefore error theory about moral beliefs is error theory about all the mental states

relevant to our moral life. Such a sweeping error theory is avoided in the present picture, since it does not touch our moral aliefs, which are also part of our moral life. Indeed, on a natural understanding of the notion of ‘practice’, moral aliefs are arguably *more* essential to the practice than moral beliefs. For the term ‘practice’ sends in the direction of the causally operative mental states in our everyday moral comportment, and in the present picture those are the aliefs, not the beliefs. If our moral aliefs form the core of our moral practice, then, and they do *not* claim any objective purport, it turns out that the present picture does not commit to error theory about the mental states at the heart of our moral practice.

Furthermore, since moral aliefs do not have an objectivistic phenomenology *qua* moral, as I argued above, they do not *present themselves* in the relevant sense as directed at objective facts. More accurately, they do not present their normative component (which is affective and/or motivational) as truth-apt. Their rational underpinnings (*qua* moral) do not depend, therefore, on the existence of moral facts and truthmakers. Rather, they depend on the aliefs being in some sense *justified*. What would make an alief justified might be its being formed in the appropriate way, or its conforming to certain standards, or something else. A story about the justification of moral aliefs is certainly needed, and a challenge could ultimately arise that would undermine in a principled manner the rational foundations of our moral alief formation. But until such a principled challenge is actually mounted, the rational underpinnings of our moral practice, being founded on traffic in moral aliefs primarily and beliefs only secondarily, remain provisionally secure.<sup>28</sup>

The primacy of alief in our moral practice can also be appreciated from a different angle, that of moral evaluation. We are routinely incensed by the self-proclaimed anti-racist when she shows traces of racist predilection in her actual behaviour. Indeed, we seem to find her more blameworthy for her racist aliefs than praiseworthy for her anti-racist beliefs. At least this is so when we evaluate *unreflective* subjects, where ‘unreflective’ means here just that the subject is not aware of any belief-alief misalignment.<sup>29</sup> A *reflective* subject may be committed to changing her racist aliefs, perhaps even be in the process of self-rehabilitation, and this appears to modify our intuitive moral evaluation of her. But as long as the subject is unreflective, so that her moral aliefs and beliefs are genuinely insulated from each other, we seem to find her more blameworthy for her racist aliefs than praiseworthy for her anti-racist beliefs. Conversely, we are charmed by the (unreflective) self-proclaimed egoist whose actual behaviour discloses deep-seated concern for others: we may find her annoying at the dinner party, but are compelled to note that ‘her heart is in the right place’. That is to say, we find her more

<sup>28</sup>As before, this is meant as an essentially conceptual claim about the notion of a practice.

<sup>29</sup>This sort of unawareness comes in two flavours. A subject may be unaware that belief sometimes mismatches behaviour, or she may be aware of the general possibility of belief-behaviour mismatch but unaware of it in her case. The first subject is, in a sense, unaware of the existence of aliefs. (To be sure, most subjects are unaware of the existence of aliefs *under that description*, but what I have in mind here is a subject unaware of the existence of underlying, unconscious mental states that can determine behaviour and may be in tension with one’s overt, conscious beliefs.) The second subject is in a sense aware of the existence of aliefs (though rarely *under that description*) but is unaware that hers misalign with her beliefs in the relevant context.

praiseworthy for her altruistic aliefs than blameworthy for her egoistic beliefs. Again, this appears to change if and when we learn that she has embarked on a deliberate process of retraining her aliefs in the goal of being a better egoist. But as long as her altruist aliefs and her egoist beliefs are genuinely insulated from each other, we tend to treat her as essentially harmless. Thus moral evaluation appears to consistently track aliefs rather than beliefs, at least when the two are insulated from each other. This is not to say that peoples' moral beliefs are of *no* concern to us, of course; merely that they are of *lesser* concern.

There is a good reason for pegging evaluation to aliefs primarily, given the design aspect discussed in §2. If in general the low-road system that puts a premium on speed over accuracy is the default system we operate on, then it is that system that evaluation should monitor for quality control in the first instance. In the domain of moral cognition, this is the alief-forming associationist mechanism.

A more ambitious way of putting the point is to say that the folk appear (for good reasons) to take a person's aliefs to constitute her *real moral commitments*. When aliefs and beliefs diverge, the *true moral character* of a person is reflected more accurately in her aliefs. (Again, at least this is so when the person is of the unreflective variety—when the divergence of belief and alief is opaque to her.) I take this to be the deep insight behind the internalist outlook in moral philosophy: real moral commitments, reflecting true moral character, are constitutively tied to the motivational states that govern her behaviour.

None of this renders moral beliefs completely irrelevant or unimportant to our moral practice. For moral beliefs, although not inherently or constitutively motivational, are not epiphenomenal either (see discussion in §1). Crucially, as we have seen above, beliefs do have the capacity to affect aliefs, if only through an Aristotelian process of self-training. At the same time, unlike aliefs, beliefs are formed through flexible personal-level processes that are consciously controllable and manipulable. In this way, our aliefs can be rationally influenced indirectly, even though they cannot be changed immediately or at will. And so a person fortunate enough to encounter Kant's second *Critique* in her early twenties might acquire the belief that she ought to treat others as ends and not mere means, and through years of careful implementation might by her late twenties or early thirties possess robust treat-as-end aliefs.<sup>30</sup> In this respect, we have no better tool for moral growth and moral progress than our moral belief-forming faculty. I take this to be the deep insight behind the cognitivist outlook in moral philosophy.

In summary, the distinction between alief and belief is serviceable not only in making sense of a range of otherwise inexplicable human behaviours, but also in unravelling what is reasonably considered the organizing problem of contemporary moral psychology: the tension between the objective purport of moral judgments and their inherently motivational

<sup>30</sup>Indeed, this appears to be the role Smith [1994] ends up assigning to moral beliefs—though I suspect he would consider the youngster much more fortunate who came across Mill's *Utilitarianism* rather than Kant's *Critique of Practical Reason*.

role. It creates the resources to articulate a conception of moral judgments ('dual-process cognitivist internalism') as dividing into two subclasses, one boasting objective purport and one inherently motivating, in a way that still captures the deep insights behind both the outlook which emphasizes objectivistic moral phenomenology and that which emphasizes constitutive moral motivation.<sup>31</sup>

University of Arizona

Received: December 2010

Revised: July 2011

## References

- Amodio, D., E. Harmon-Jones, and P. Devine 2003. Individual Differences in the Activation and Control of Affective Race Bias as Assessed by Startle Eyeblink Response and Self-Report. *Journal of Personality and Social Psychology* 84/4: 738–53.
- Armstrong, D. M. 1968. *A Materialist Theory of the Mind*, New York: Humanities Press.
- Arpaly, Nomy 2002. Moral Worth, *Journal of Philosophy* 99/5: 223–45.
- Ayer, A. J. 1936. *Language, Truth, and Logic*, London: Victor Gollancz.
- Barrett, L. F., M. M. Tugade, and R. W. Engle 2004. Individual Differences in Working Memory Capacity and Dual-Process Theories of the Mind, *Psychological Bulletin* 130/4: 553–73.
- Brink, David 1989. *Moral Realism and the Foundations of Ethics*, Cambridge: Cambridge University Press.
- Cohen, L. Jonathan. 1992. *An Essay on Belief and Acceptance*, Oxford: Clarendon Press.
- Copp, D. 2001. Realist-Expressivism: A Neglected Option for Moral Realism, *Social Philosophy and Policy* 18/2: 1–43.
- Falk, W. D. 1945. Obligation and Rightness, *Philosophy* 20/76: 129–47.
- Frankena, W. K. 1958. Obligation and Motivation in Recent Moral Philosophy, in *Essays in Moral Philosophy*, ed. Abraham Melden, Seattle: University of Washington Press: 40–81.
- Geach, P. T. 1960. Ascriptivism, *Philosophical Review* 69/2: 221–25.
- Gendler, T. S. 2008a. Alief and Belief, *Journal of Philosophy* 105/10: 634–63.
- Gendler, T. S. 2008b. Alief in Action (and Reaction), *Mind and Language* 23/5: 552–85.
- Gibbard, Allan 2003. *Thinking How to Live*, Cambridge, MA: Harvard University Press.
- Greene, Joshua and Jonathan Haidt 2002. How (and Where) Does Moral Judgment Work?, *Trends in Cognitive Science* 6/12: 517–23.
- Hare, R. M. 1952. *The Language of Morals*, Oxford: Clarendon Press.
- Horgan, T. and J. Tienson 2002. The Intentionality of Phenomenology and the Phenomenology of Intentionality, in *Philosophy of Mind: Classical and Contemporary Readings*, ed. David Chalmers, Oxford and New York: Oxford University Press: 520–33.
- Horgan, T. and M. C. Timmons, 2005. Moral Phenomenology and Moral Theory, *Philosophical Issues* 15/1: 56–77.
- Horgan, T. and M. C. Timmons 2006. Cognitivist Expressivism, in *Metaethics After Moore*, ed. Terry Horgan and Mark Timmons, Oxford: Oxford University Press: 255–98.
- Horgan, T. and M. C. Timmons 2007. Moorean Moral Phenomenology, in *Themes from G. E. Moore*, ed. S. Nuccetelli and G. Seay, Oxford: Oxford University Press: 203–26.
- Horgan, T. and M. C. Timmons 2008. Prolegomena for a Future Phenomenology of Morals, *Phenomenology and the Cognitive Sciences* 7/1: 115–31.
- Kahneman D. 2003. A Perspective on Judgement and Choice, *American Psychologist* 58: 697–720.
- Kalderon, Mark Eli 2005. *Moral Fictionalism*, Oxford: Oxford University Press.
- Kriegel, Uriah 2003. Consciousness as Sensory Quality and as Implicit Self-Awareness, *Phenomenology and the Cognitive Sciences* 2/1: 1–26.
- Kriegel, Uriah, ed, 2008. *Moral Phenomenology. Special issue of Phenomenology and the Cognitive Sciences* 7/1.
- Kriegel, Uriah 2008a. Moral Phenomenology: Foundational Issues, *Phenomenology and the Cognitive Sciences* 7/1: 1–19.
- Kriegel, Uriah forthcoming. *The Varieties of Consciousness: Essays on Non-Sensuous Phenomenology*, Oxford and New York: Oxford University Press.
- McDowell, John 1979. Virtue and Reason, *The Monist* 62/3: 331–50.
- McNaughton, David 1988. *Moral Vision*, Oxford: Blackwell.
- Mackie, J. L. 1977. *Ethics: Inventing Right and Wrong*, New York: Penguin.
- Mandelbaum, Maurice 1955. *The Phenomenology of Moral Experience*, Glencoe IL: The Free Press.
- Milner, David and Melvin Goodale 1995. *The Visual Brain in Action*, Oxford: Oxford University Press.
- Moore, G. E. 1953. Propositions, in his *Some Main Problems of Philosophy*, London: George Allen and Unwin: 52–71.

<sup>31</sup>For comments on a pair of previous drafts, I am greatly indebted to Tamar Gendler, Mark Timmons, and an outstanding anonymous referee for this journal.

- Peacocke, C. 1998. Conscious Attitudes, Attention, and Self-Knowledge, in *Knowing Our Own Minds*, ed. Crispin Wright, Barry Smith, and Cynthia Macdonald, Oxford: Oxford University Press: 63–98.
- Pitt, D. 2004. The Phenomenology of Cognition; or What Is It Like to Think that P?, *Philosophy and Phenomenological Research* 69/1: 1–36.
- Rosenthal, D. M. 1990 (1997). A Theory of Consciousness, ZiF Technical Report 40, Bielefeld, Germany, reprinted in *The Nature of Consciousness: Philosophical Debates*, ed. Ned Block, Owen Flanagan, and Güven Güzeldere, Cambridge MA: MIT Press: 729–54.
- Scheler, Max 1913 (1973). *Formalism in Ethics and Non-Formal Ethics of Values*, trans. M. S. Frings and R. L. Funk, Evanston IL: Northwestern University Press.
- Siewert, Charles 1998. *The Significance of Consciousness*, Princeton NJ: Princeton University Press.
- Simons, Peter 1987. *Parts: A Study in Ontology*, Oxford: Clarendon Press.
- Sloman, S. A. 1996. The Empirical Case for Two Systems of Reasoning, *Psychological Bulletin* 119/1: 3–22.
- Smith, Michael 1994. *The Moral Problem*, Oxford: Blackwell.
- Strawson, Galen 1994. *Mental Reality*, Cambridge MA: MIT Press.
- Sun, R., P. Slusarz, and C. Terry 2005. The Interaction of the Explicit and the Implicit in Skill Learning: A Dual-Process Approach, *Psychological Review* 112/1: 159–92.

815

820

825

830

835

840

845

850

855

860