

Patients With Ventromedial Frontal Damage Have Moral Beliefs

Adina Roskies

Michael Cholbi thinks that the claim that motive internalism (MI), the thesis that moral beliefs or judgments are intrinsically motivating, is the best explanation for why moral beliefs are usually accompanied by moral motivation. He contests arguments that patients with ventromedial (VM) frontal brain damage are counterexamples to MI by denying that they have moral beliefs. I argue that none of the arguments he offers to support this contention are viable. First, I argue that given Cholbi's own commitments, he cannot account for VM patients' behavior without attributing moral beliefs to them. Secondly, I show that his arguments that we should not believe their self-reports are unconvincing. In particular, his argument that they cannot self-attribute moral beliefs because they have a defective theory of mind is flawed, for it relies upon a misreading of both the empirical and theoretical literatures. The avenues remaining to Cholbi to support motive internalism are circular, for they rely upon an internalist premise. I provide an alternative picture consistent with neuroscientific and psychological data from both normals and those with VM damage, in which connections between moral belief and motivation are contingent. The best explanation for all the data is thus one in which MI is false.

Keywords: Moral Belief; Moral Judgment; Motive Internalism; Theory Of Mind; Ventromedial Frontal Damage

In "Belief Attribution and the Falsification of Motive Internalism," Michael Cholbi (2006) raises interesting and novel objections to taking patients with ventromedial (VM) frontal damage as challenges to motive internalism (MI) in ethics. Rather than adopting standard arguments against empirical assaults on internalism, such as maintaining that the people in question lack moral concepts (see Kennett & Fine, in

Correspondence to: Adina Roskies, Department of Philosophy, 207B Thornton Hall, Dartmouth College, Hanover, NH 03755, USA. Email: adina.roskies@dartmouth.edu

press; Smith, 1994), that they make moral judgments only in an “inverted commas” sense (see Hare, 1956; Kennett & Fine, in press; Smith, 1994), or that they are in fact motivated (Kennett & Fine, in press), Cholbi makes two novel claims. First, that despite their intact reasoning ability and knowledge, there are reasons to doubt that these people in fact possess the relevant moral beliefs, and second, that their inability to attribute moral beliefs to others impairs their ability to attribute them to themselves. Cholbi further maintains that in light of these failings, internalism is the *prima facie* most plausible explanation of the correspondence often or normally found between moral belief and moral motivation.

Cholbi and I agree on a number of crucial points about VM patients. To wit: VM patients make moral judgments, they fail to be motivated by them, they have unimpaired moral reasoning, and finally, their moral concepts are intact. Despite these substantive points of agreement, Cholbi maintains that we have reason to doubt that VM patients’ moral judgments provide evidence that they actually hold moral beliefs.

Although Cholbi accepts that VM patients have unimpaired moral reasoning abilities, he argues that competence in moral reasoning is not evidence for belief possession. In particular, he claims that moral reasoning requires an ability to reason hypothetically about morality, but that the ability of VM patients to so reason provides insufficient evidence that VM patients actually believe the antecedents of the conditionals with which they reason:

But competence in moral reasoning relative to a body of propositions is no evidence of an agent’s believing those propositions, and this need not be explained by the subject’s failure to understand the relevant concepts. . . . Proficiency in moral reasoning only requires belief in conditional moral propositions . . . irrespective of whether the antecedents of such propositions are believed. (2006, p. 610)

Cholbi instead explains the ability of VM patients to reason normally as a general ability to reason counterfactually, and he likens them to psychopaths, who seem to lack the ability to distinguish moral rules from conventional rules. He argues that denying that VM patients hold moral beliefs is at least as plausible an account of their lack of motivation as is the denial that their beliefs are motivationally efficacious.

As far as I know, VM patients are not relevantly like psychopaths, since they do not seem to lack the ability to make the moral–conventional distinction (see, e.g., Saver & Damasio, 1991). But let us consider the possibility that Cholbi raises: that VM patients reason hypothetically but do not believe the antecedents of the conditionals they use. Although one could imagine that such a profile could account for the finding that their reasoning is unimpaired, it could not account for the finding that they answer moral questions the way that normals do. For if the propositions they used to reason were all hypothetical, such as “if enslaving human beings is morally wrong, then so is selling them,” then they would not be able to provide the answer that normals give to the question “Is it wrong to sell human beings?” That answer will depend upon which antecedent proposition one begins with: if someone holds “Enslaving humans is wrong” she will answer “yes”; if she holds “Enslaving humans

is not wrong” she will answer “no.” If she holds neither, she cannot answer the question. In short, to answer any question of the form, “Is it morally right (or wrong) to . . .” one must employ a nonhypothetical moral proposition. It therefore cannot be the case that VM patients provide answers normals provide merely by hypothetical reasoning based solely on conditional statements.

Cholbi might still maintain that VM patients hold nonhypothetical propositions to be true but that this attitude of holding fails to be one of belief. He suggests instead the attitude is one of supposition. But since VM patients give the same answers to moral questions that normals do, it must be that the VM patients only suppose the propositions that normals believe. Now the burden is on Cholbi to explain why we should deny that these suppositions are beliefs, and further, to do so without begging the question by relying upon the truth of internalism. One obvious possibility is that they report moral statements they think others hold, so that they make moral judgments in the “inverted commas” sense. But Cholbi agrees with me that this is not the right interpretation of their behavior: “I am not claiming . . . that VM patients are utilizing moral language in an ‘inverted commas’ sense” (2006, p. 613).

It is unclear whether Cholbi’s point is that VM patients fail to hold any moral beliefs, or just that they don’t hold the moral beliefs that they appear to affirm in experimental situations, and are unmoved by. Let us consider both possibilities. If the former, then, as discussed above, it looks as if Cholbi cannot explain why or how their moral reasoning results in moral statements with which normal people concur. If the latter, Cholbi must agree that VM patients do have moral beliefs, but not the ones that they employ in reasoning. He then owes us an account of why it is that the moral beliefs the VM patients hold are not the ones they employ in reasoning. His account cannot involve taking their claims to be merely reports of conventional norms, since he does not think they make moral judgments in the “inverted commas” sense. But even if we allowed them moral beliefs in the inverted commas sense, if they do have real moral beliefs but don’t use them in the moral reasoning explored in experimental situations, then presumably the moral beliefs they do hold are not the same as the moral suppositions they employ in reasoning. We then need an account of (a) why they employ these suppositions instead of their true moral beliefs, (b) why damage to VM cortex leads to a disruption of belief such that the normal beliefs they held prior to injury are either abolished or supplanted by different or even conflicting beliefs which do not figure in reasoning, and (c) while the content of moral beliefs prior to injury is then transferred to mental states that do not count as belief, but are the mental states used in moral reasoning. While Cholbi may address (a) by reference to the artificial stress of experimental situations (see below), explaining (b) and (c) seem like difficult tasks indeed. They are made especially difficult when we consider the evidence from VM patients upon which we both agree.

The evidence we both agree upon is that VM patients assent to moral propositions when questioned, they do not appear to try to deceive their questioner (Cholbi, 2006, p. 613), and no account of the effects of damage to VM cortex suggests that prior knowledge is disrupted—indeed, Cholbi accepts that their understanding

of moral concepts is intact: "I am not claiming...that VM patients fail to understand the moral concepts contained in their moral beliefs" (p. 613). However, if understanding moral concepts is moral knowledge, since knowledge is (with the appropriate caveats) justified belief, that entails that at least some moral beliefs are intact. Moreover, Cholbi does not offer a systematic explanation that provides good reason to doubt that VM patients have moral beliefs and at the same time accounts for the evidence that appears to point strongly to them having the moral beliefs they appear to have. The most plausible account of the data is simply that VM patients believe nonhypothetical moral statements.

I believe that this argument is sufficient to refute Cholbi's claim that competence in moral reasoning (hypothetical) is not evidence for possession of belief, but let me provide a brief sketch of a positive view to further bolster the claim that VM patients have moral beliefs (see also Roskies, *in press*). My working model of how VM cortex is involved in moral belief and motivation is that VM cortex is necessary for acquisition of moral concepts, but not their retention or employment. Damage to VM cortex results in disconnection of the pathway by which cognitive processing of moral propositions normally causes activation of emotional and motivational systems that ultimately lead to action. This model explains why moral reasoning usually results in moral motivation, why damage to VM cortex in early life prevents people from learning moral concepts (Anderson, Bechara, Damasio, Tranel, & Damasio, 1999), and why the connection between moral belief and motivation is contingent and not necessary, and thus why the form of MI I target is false. It also explains why damage to VM cortex (of the type discussed in Roskies, 2003) fails to impair the moral concepts and beliefs of VM patients who have already acquired moral knowledge. Just as patients with damage to hippocampus are impaired in acquiring new declarative memories but retain their pre-trauma declarative memory (Milner, 2005), VM patients are impaired in acquiring new moral knowledge but not in retaining or employing moral beliefs. We are not tempted to deny that a patient with hippocampal damage knows who he is or that he really believes the pre-trauma memories which he recounts and reasons about. Similarly, we should not be tempted to deny that patients with VM damage believe the moral propositions they endorse and employ in reasoning.

In Cholbi's (2006) discussion of whether there is sufficient evidence to attribute moral beliefs to VM patients, he raises questions about whether the linguistic evidence I rely upon in Roskies (2003) is reliable evidence for moral belief. He points to (a) "standard worries that the biases or pressures introduced by the experimental context" may influence the sincerity of the VM patients' responses (p. 611), suggesting that experiments about moral reasoning are likely to elicit conventional responses even in the absence of belief. He also suggests that (b) since belief is not subject to the law of excluded middle, VM patients will be more likely to err or mislead in their self-reports (p. 611). My response to both these claims is that these worries attend experiments with any subjects, including normals, and without independent reason to think that VM patients will be more subject to these influences than normals, or that they in fact lack moral beliefs, we should treat their responses

on a par with those of normals. Since Cholbi offers no reason to think differently of VM patients and normals, I am no more inclined to believe that VM responses are less reliable than the responses of normals. We certainly accord moral beliefs to normal experimental subjects, and rely upon their linguistic reports, despite the possibility that experimental situations compromise the veracity of those reports. The fact that VM patients fail to act in socially appropriate ways in both normal and experimental contexts (Bechara, Damasio, & Damasio, 2000; Damasio, Tranel, & Damasio, 1990; Eslinger & Damasio, 1985; Saver & Damasio, 1991) provides additional reason to deny Cholbi's suggestion that VM patients are moved to behave (linguistically or otherwise) in conventional ways in experimental situations. If one is worried about appearing to abide by strong social norms, one would be more apt to be careful in exhibiting appropriate moral behavior in experimental contexts than merely in uttering appropriate moral sentences (which, in the absence of belief, is making moral judgments in the "inverted commas" sense, an explanation that Cholbi disavows).

Cholbi (2006) next argues that even given the linguistic evidence, it is "arguable that it is defeated or outweighed by the body of contrary affective and behavioral evidence" (p. 612). Here he appeals to the fact that beliefs have a causal role in the production of behavior, so that "an agent does not believe that *P* unless she behaves as a person who believes that *P* characteristically behaves, where the behavior in question is not only linguistic" (p. 612). I take it that Cholbi thinks that we lack reason to take the VM patients' linguistic reports as evidence for moral beliefs in light of the fact that they fail to act accordingly, and fail to exhibit the affective responses normals would in similar situations. He rightly points out that all we can appeal to as evidence for belief in the case of nonlinguistic creatures is their behavior, and that linguistic evidence in the case of humans is simply additional, but not authoritative, evidence (p. 612). Although I agree that nonlinguistic behavior plays an important role in attributing beliefs to both people and animals, we often take linguistic report to be privileged with respect to certain questions, among them questions of belief (see, e.g., Davidson, 1975/1984; Malcolm, 1972). To deny an agent moral beliefs on the basis of his nonlinguistic behavior alone would presumably be on the basis of an argument like the following (call it "CB"), which I believe is the argument Cholbi has in mind:

- CB1. An agent does not believe that *P* unless she behaves as an agent who believes that *P* characteristically behaves.
- CB2. An agent that believes that it is right to Φ in *C* is characteristically motivated Φ to in *C*, to do *A* in *C*, and to express emotion *E* in *C*.
- CB3. If an agent is not motivated to Φ in *C*, to do *A* in *C*, and to express emotion *E* in *C*, she does not believe that it is right to Φ in *C*.

(CB2) looks much like the strong internalist claim (SI) that I use VM patients to target: "If an agent believes that it is right to Φ in circumstances *C*, then he is motivated to Φ in *C*" (Roskies, 2003, p. 55). The differences here are (a) that the internalist claim is stronger than Cholbi's, for the internalist holds that the connection between moral belief and motivation is necessary, while Cholbi suggests that it is merely characteristic, and (b) Cholbi appeals to emotion and action as

evidence additional to motivation. However, looking more closely it is difficult to see how this argument works.

First, let us consider whether the additional evidence that Cholbi appeals to makes his case better than merely appealing to lack of motivation. On the reasonable assumptions, consistent with neuroanatomy and my model, that VM cortex is upstream of (and causally connected to) central emotional and motivational structures (e.g., the limbic system), and that action requires motivation, it does not seem that appealing to emotion and action in fact marshals more independent evidence than appealing to motivation alone. That is, VM damage would be expected to influence emotion, motivation, and action. Thus we can simplify premise (CB2) to read:

CB2'. An agent that believes that it is right to Φ in C is characteristically motivated to F in C .

In (CB1), (CB2) and (CB2') the term 'characteristically' either plays no substantive role, or it does. If the former, (CB2') is equivalent to "An agent that believes that it is right Φ in C is motivated to Φ in C ", which is simply thesis (SI), so a premise in Cholbi's argument for why to discount VM data is equivalent to precisely the thesis which I have employed VM data to dispute. If this is Cholbi's argument, it begs the question. On the other hand, if 'characteristically' is to play a substantive role, what is that role? Since people who claim to believe P can behave in innumerable different ways (even behave as if $\sim P$, if they are lying, forgetful, etc.) it seems impossible to define characteristic behavior for believing that P , for any P . It is therefore impossible to use (CB1) as a necessary criterion for belief attribution. Furthermore, if we follow the spirit of Cholbi's argument and merely interpret 'characteristically' as a claim about most but not all people's general pattern of behavior, then it is clear that there are cases in which an agent believes that P but does not behave characteristically. I claim that VM patients present such a case, and my sketch of the functional organization of behavior above suggests why that is. Moreover, given that interpretation of 'characteristically', (CB1) is false and argument CB is unsound. So we cannot conclude that VM patients do not have moral beliefs. That is a good thing, for if argument CB were true, either there would be no examples of practical irrationality, for beliefs would always mesh with behavior, or in cases where they didn't, everyone would be practically irrational.

Cholbi (2006) concludes this part of his argument thus: "Why, even if we suppose that linguistic affirmations of belief are evidence at all, should that evidence be thought of as decisive when it contradicts other affective and behavioral evidence?" (p. 612). I do not claim that the evidence is decisive merely because it is linguistic evidence, but in my 2003 paper and in my above sketch of the functional organization of VM cortex and regions subserving motivation and affect, I offer a number of independent reasons to think that beliefs remain intact with VM damage, and that VM patients' linguistic behavior is as good a window onto their beliefs as it is in the case of normals. I will not reiterate those arguments here.

Cholbi's most interesting challenge to taking VM patients as counterexamples to internalism is his argument (let's call it "BA") that: (1) VM patients have difficulty interpreting others' beliefs and behaviors; (2) that there is a common neural mechanism for self-attribution of beliefs and attribution to others; and (3) that VM patients lack an "adequate theory of mind." The conclusion is (4) that we have reason to doubt VM patients' belief self-reports.

One need not assent to the Cartesian view that one always has incorrigible access to one's beliefs to think that one has reliable access to many of them. Even if one accepts the Davidsonian hypothesis that to be a believer one has to be an interpreter of beliefs, one need not hold that being a believer requires particularly insightful abilities in the belief-attribution-to-others department. Consider Davidson's views; I choose Davidson because he posits unusually strict criteria for being a believer, so strict that he denies nonhuman animals beliefs. Despite his extreme views, Davidson's criteria involve only general abilities, not specific ones on a proposition-by-proposition basis. For instance, in various places Davidson writes that all you need to be a believer is to be a speaker of a language (Davidson, 1975/1984), to have a concept of belief (Davidson, 1975/1984, 2001) or to have a conception of truth and the possibility of misrepresentation (Davidson, 1975/1984, 2001). VM patients retain all these concepts and abilities. So even if self-attribution of belief is interdependent with the ability to attribute beliefs to others, this does not entail that this ability to attribute beliefs to others must be flawless or even very good. In laying out argument BA, Cholbi (2006) writes:

1. It is possible for individual S_1 to justifiably attribute to herself the belief that P only if S_1 can justifiably attribute P to other individuals S_2, \dots, S_n . (p. 613)

This cannot possibly be true. Fermat presumably attributed to himself the belief that "It is impossible for a cube to be the sum of two cubes, a fourth power to be the sum of two fourth powers, or in general for any number that is a power greater than the second to be the sum of two like powers," but had he attributed such a belief to anyone else, he would have been unjustified in doing so. It surely must be possible to attribute a belief to oneself that one cannot justifiably attribute to others, for there must be a possibility that someone has a new and unique belief, and recognizes it as such. It may be the case that for any belief B , for one to be able to self-attribute B it must be possible for one to conceive of attributing B to someone else, but this is not equivalent to (1). Indeed, in the literature on self-ascription, it is general capacities that matter, not specific beliefs or specific types of beliefs.

Cholbi's (2006) argument BA continues:

2. VM patients cannot justifiably attribute moral beliefs to other subjects.
3. It is not possible for VM patients to justifiably self-attribute moral beliefs.
[therefore]
4. VM patients' linguistic reports of moral belief are *prima facie* unreliable. (p. 613)

Regarding (2), it is not clear that this is correct. What is known is that VM cortex is involved in emotional experience and in the recognition of emotion in others

(Ochsner, Knierim, & Ludlow, 2004; Shamay-Tsoory, Tomer, & Berger, 2005; Shamay-Tsoory, Tomer, Berger, & Aharon-Peretz, 2003). Damage to VM cortex results in difficulties in attributing emotional states to others on the basis of facial and vocal characteristics (Shamay-Tsoory et al., 2003), and leads to the disruption of the subjective experience of emotion, as indicated by self-report (Bechara et al., 2000, Damasio et al. 1990). What we can conclude from these studies is that VM patients have emotional deficits, and have difficulty in attributing emotions to others, and thus that they may not be reliable in emotion attribution. It is a further and quite substantive step to conclude that VM patients cannot attribute moral beliefs to others, or are not justified in doing so. While brain regions involved in emotional processing are sometimes involved in moral judgment (Greene & Haidt, 2002; Greene, Sommerville, Nystrom, Darley, & Cohen, 2001), it is likely that not all moral reasoning relies upon these areas, so it appears to be false that one must be emotionally intact in order to have moral beliefs. Furthermore, it is not at all clear that we normally attribute moral beliefs to others on the basis of emotional expressions, let alone that we need to in order to be justified. Indeed, our primary access to the moral beliefs of others is their linguistic affirmations and/or the recognition of generally accepted moral statements. There is no evidence that VM patients are impaired on either of these matters.

Cholbi's argument seems to rest upon the idea that having a theory of mind (ToM) is necessary for proper attribution of beliefs to others or oneself. It is often argued that ToM is important for belief attribution (Frith & Frith, 2003; Saxe, Carey, & Kanwisher, 2004; Wimmer & Perner, 1983; but see Nichols & Stich, 2003, for an argument that self-attribution of mental states is independent of a theory of mind). As Cholbi reports, Shamay-Tsoory et al. (2003, 2005) claim that VM patients are impaired in a ToM test. However, the version of ToM tested in the Shamay-Tsoory papers is not the standard ToM test administered in the literature, which requires only an understanding of false belief (Saxe et al., 2004; Wimmer & Perner, 1983). In these classic studies, the premise is that having a rich concept of belief is necessary for understanding what beliefs are, and thus for proper belief ascription. However, in study cited by Cholbi, VM patients only showed deficits in "affective theory of mind" as assayed by responses to a faux-pas assay, which requires both an understanding of other's beliefs and an ability to assess the effect upon their emotional states (Shamay-Tsoory et al., 2003, 2005). Given the emotional deficits VM patients suffer (see above), it is unsurprising that VM patients are deficient in this test, for in addition to standard ToM requirements, proper answers rely upon emotional processing. In fact, in this same study, Shamay-Tsoory et al. (2003) reports that VM patients showed *no* deficit in ordinary false-belief tasks. Cholbi (2006) himself concurs that "there is no clear basis for concluding that VM patients are conceptually incompetent" (p. 613). There is, therefore, no evidence that their concept of belief is at all impaired, or that they have difficulty in attributing beliefs to others. Therefore, even if we accept that having a theory of mind is important for attributing beliefs to others, we lack reason to think that VM patients are unable to reliably attribute beliefs, for they are not impaired on standard theory of mind tasks.

Even if VM patients were impaired in correctly attributing beliefs to others, to conclude that reliable self-attribution of belief requires reliable attribution of belief to others needs further argument. To shore up this claim, Cholbi makes reference to studies that show that VM cortex is activated both in self-attribution of emotional states and other-attribution of these same states (Oschner et al., 2003). First, this evidence pertains to attributions of emotional states, and not belief, let alone moral belief. Second, even if it is the case that some of the same brain regions are activated in both types of operations it does not follow that damage to some of these areas undermines the ability to reliably attribute beliefs to self or other. This requires further evidence, ideally based on localized brain lesions. There are many redundancies in brain organization, and there is probably more than one neural pathway that can be used for self-attribution of mental states. Notice, for example, that the evidence that Cholbi and I rely upon to support the claim that the experience of emotion is impaired in VM patients comes from self-report. Unless we want to deny this as relevant evidence about the syndrome of VM damage, we must acknowledge some ability to self-attribute emotional states, even when the ability to attribute such states to others is defective.

There are independent reasons to deny Cholbi the move from (2) to (3). It is clear that the access we have to others' emotional or mental states is different than the access we have to our own. We assess the emotional state of others by interpreting their facial expressions, their tone of voice, and their words and actions, but we do not use these same means, at least not to the same extent, in figuring out our own emotional states. Therefore it is at least conceivable that one could be impaired at recognizing mental states in others but not oneself. Finally, it is important to remember that VM patients already possess moral and emotional knowledge. Even if recognition of mental states is impaired in others, self-recognition of similar states might be intact. So even though processing social or emotional stimuli is impaired in VM patients, we need not conclude that they are impaired in reporting their own mental states.

In summary, premise (1) of BA is false, and if reconstructed to be plausible, it is not clear that VM patients run afoul of it. There is inadequate evidence for (2), for it rests upon taking evidence for failures in emotional attribution to be paralleled by failures in moral attribution. The move from (2) to (3) is unwarranted, since it rests on an assumption that the ToM assay in Shamay-Tsoory et al. (2003, 2005) is equivalent to standard ToM tests, and because there are plausibly different cognitive routes for self-ascription of belief than the route arguably impaired by VM cortex damage. I think that (3) is likely false, though I know of no direct tests of it. Therefore, I don't think there is reason to accept (4). Cholbi (2006) himself recognizes that BA is not demonstrative, for he admits that (3) "is too coarse" (p. 613) and that the argument "does not demonstrate that the moral beliefs in question are absent," but rather suggests that we should not afford them much evidentiary weight (p. 613).

Cholbi (2006) concludes that MI is subject to empirical challenge, but that VM patients do not provide a compelling challenge, for he rejects "the *prima facie* evidence

for imputing motivationally inert beliefs to VM patients” (p. 613). I agree that

A comprehensive philosophical psychology of morality should aim to explain as wide a body of phenomena as possible, and one fact that both motive internalists and their opponents accept (and therefore bear the burden of explaining) is that moral beliefs often do motivate agents to act. (p. 614)

However, I strongly disagree with what follows. Cholbi proposes:

That the best defense of MI is to see it as the best empirical explanation of this fact... We should therefore treat moral belief as having a distinctively motivating role unless there is evidence that shows that MI assigns the wrong role to moral belief in our cognitive architecture... the evidence on which the attribution of moral belief is based seems to count as evidence only given the presumption that moral beliefs are intrinsically motivating states. (p. 614)

I counter instead that the best empirical explanation of a wide variety of phenomena, including cases of brain damage and common observations that moral beliefs often do motivate is my very plausible picture of the role that VM cortex plays in linking beliefs with emotional and motivational brain systems. On that picture, moral beliefs are causally connected with motivation in normal people because of how the brain is wired, but causal connections are contingent, and so moral beliefs are not intrinsically motivating.

References

- Anderson, S. W., Bechara, A., Damasio, H., Tranel, D., & Damasio, A. R. (1999). Impairment of social and moral behavior related to early damage in human prefrontal cortex. *Nature Neuroscience*, 2, 1032–1037.
- Bechara, A., Damasio, H., & Damasio, A. R. (2000). Emotion, decision making and the orbitofrontal cortex. *Cerebral Cortex*, 10, 295–307.
- Cholbi, M. J. (2006). Belief attribution and the falsification of motive internalism. *Philosophical Psychology*, 19, 607–616.
- Damasio, A. R., Tranel, D., & Damasio, H. (1990). Individuals with sociopathic behavior caused by frontal damage fail to respond autonomically to social stimuli. *Behavioural Brain Research*, 41, 81–94.
- Davidson, D. (1984). Thought and talk. In *Inquiries into truth and interpretation* (1st ed., pp. 155–170). Oxford, England: Clarendon Press. (Original work published 1975)
- Davidson, D. (2001). Rational animals. In *Subjective, intersubjective, objective* (pp. 95–106). Oxford, England: Clarendon Press. (Original work published 1982)
- Eslinger, P. J., & Damasio, A. R. (1985). Severe disturbance of higher cognition after bilateral frontal lobe ablation: Patient EVR. *Neurology*, 35, 1731–1741.
- Frith, U., & Frith, C. D. (2003). Development and neurophysiology of mentalizing. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 358, 459–473.
- Greene, J. D., & Haidt, J. (2002). How (and where) does moral judgment work? *Trends in Cognitive Sciences*, 6, 517–523.
- Greene, J. D., Sommerville, R. B., Nystrom, L. E., Darley, J. M., & Cohen, J. D. (2001). An fMRI investigation of emotional engagement in moral judgment. *Science*, 293, 2105–2108.
- Hare, R. M. (1956). *The language of morals*. Oxford, England: Clarendon Press.

- Kennett, J., & Fine, C. (in press). Internalism and the evidence from psychopaths and “acquired sociopaths.” In W. Sinnott-Armstrong (Ed.), *Moral psychology: Vol. 3: The neuroscience of morality*. Cambridge, MA: MIT Press.
- Malcolm, N. (1972). Thoughtless brutes. *Proceedings and Addresses of the American Philosophical Association*, 46, 5–20.
- Milner, B. (2005). The medial temporal-lobe amnesic syndrome. *Psychiatric Clinics of North America*, 28, 599–611.
- Nichols, S., & Stich, S. (2003). *Mindreading*. Oxford, England: Oxford University Press.
- Oschner, K. N., Knierim, K., & Ludlow, D. H. (2004). Reflecting upon feelings: An fMRI study of neural systems supporting the attribution of emotion to self and other. *Journal of Cognitive Neuroscience*, 16, 1746–1772.
- Roskies, A. (2003). Are ethical judgments intrinsically motivational? Lessons from “acquired sociopathy.” *Philosophical Psychology*, 16, 51–66.
- Roskies, A. L. (in press). Internalism and the evidence from pathology. In W. Sinnott-Armstrong (Ed.), *Moral psychology: Vol. 3: The neuroscience of morality*. Cambridge, MA: MIT Press.
- Saver, J. L., & Damasio, A. R. (1991). Preserved access and processing of social knowledge in a patient with acquired sociopathy due to ventromedial frontal damage. *Neuropsychologia*, 29, 1241–1249.
- Saxe, R., Carey, S., & Kanwisher, N. (2004). Understanding other minds: Linking developmental psychology and functional neuroimaging. *Annual Review of Psychology*, 55, 87–124.
- Shamay-Tsoory, S. G., Tomer, R., & Berger, B. D. (2005). Impaired “affective theory of mind” is associated with right ventromedial prefrontal damage. *Cognitive and Behavioral Neurology*, 18, 55–67.
- Shamay-Tsoory, S. G., Tomer, R., Berger, B. D., & Aharon-Peretz, J. (2003). Characterization of empathy deficits following prefrontal brain damage: The role of the right ventromedial prefrontal cortex. *Journal of Cognitive Neuroscience*, 15, 324–337.
- Smith, M. (1994). *The moral problem*. Oxford, England: Blackwell.
- Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children’s understanding of deception. *Cognition*, 13, 103–128.