

# Morphological Rationalism and the Psychology of Moral Judgment

Terry Horgan · Mark Timmons

Accepted: 16 March 2007 / Published online: 2 May 2007  
© Springer Science + Business Media B.V. 2007

**Abstract** According to rationalism regarding the psychology of moral judgment, people's moral judgments are generally the result of a process of reasoning that relies on moral principles or rules. By contrast, intuitionist models of moral judgment hold that people generally come to have moral judgments about particular cases on the basis of gut-level, emotion-driven intuition, and do so without reliance on reasoning and hence without reliance on moral principles. In recent years the intuitionist model has been forcefully defended by Jonathan Haidt. One important implication of Haidt's model is that in giving reasons for their moral judgments people tend to confabulate – the reasons they give in attempting to explain their moral judgments are not really operative in producing those judgments. Moral reason-giving on Haidt's view is generally a matter of post hoc confabulation. Against Haidt, we argue for a version of rationalism that we call 'morphological rationalism.' We label our version 'morphological' because according to it, the information contained in moral principles is embodied in the standing structure of a typical individual's cognitive system, and this morphologically embodied information plays a causal role in the generation of particular moral judgments. The manner in which the principles play this role is via 'proceduralization' – such principles operate automatically. In contrast to Haidt's intuitionism, then, our view does not imply that people's moral reasoning practices are matters of confabulation. In defense of our view, we appeal to what we call the 'nonjarring' character of the phenomenology of making moral judgments and of giving reasons for those judgments.

**Keywords** Confabulation · Intuitionism in moral psychology · Moral experience · Moral judgment · Moral phenomenology · Morphological rationalism · Rationalism in moral psychology

---

T. Horgan (✉) · M. Timmons  
Department of Philosophy, Social Sciences 213, University of Arizona, Tucson, AZ, USA  
e-mail: thorgan@u.arizona.edu

M. Timmons  
e-mail: mtimmons@u.arizona.edu

In recent years, moral psychologists and empirically-minded philosophers have challenged the idea that moral principles play an important role in the psychological economy of agents. The focus of this recent work is on peoples' everyday moral judgments about particular cases, and how they come to have or make those judgments. According to rationalist models of moral judgment, of the sort championed by Piaget (1932) and Kohlberg (1981), peoples' moral judgments are generally the result of a process of reasoning that relies on moral principles or rules. By contrast, intuitionist models of moral judgment – the new trend in moral psychology – hold that people generally come to have moral judgments about particular cases on the basis of gut-level, emotion-driven intuition, and do so without reliance on reasoning and hence without reliance on moral principles. On the intuitionist view, then, moral principles and rules are 'marginalized' – they are far less important in the explanation of people's moral judgments than on rationalist models.

We think the rationalist/intuitionist opposition, as currently understood in the field of moral psychology, fails to recognize certain options regarding the psychology of moral judgment. In particular, it fails to recognize an option that we think is not only intuitively very plausible, but fits better with certain aspects of people's everyday moral experience – their moral phenomenology – than do opposing views. We call our view 'morphological rationalism' (MR, for short). It is a version of rationalism because it sides with standard forms of rationalism in affirming an important and perhaps ineliminable role for moral principles in people's moral judgments. We label our version 'morphological' because according to it, the information contained in moral principles is embodied in the standing structure of a typical individual's cognitive system, and this morphologically embodied information plays a causal role in the generation of particular moral judgments. The manner in which the principles play this role is via 'proceduralization' – such principles operate automatically. To be guided by a moral principle, one need not think about or otherwise 'call up' the principle. On the view we propose, then, moral principles do not need to be occurrently 'represented' (either consciously or unconsciously) by the system in order to be playing the relevant causal role. Finally, our view embraces the basic intuitionist idea that in a great many cases, people's moral judgments are psychologically spontaneous, and so not the result of conscious (or even unconscious) reasoning.

Explaining and then motivating this option are our primary aims in this paper. MR about moral judgment is intended as an empirical hypothesis. As philosophers, we hope to contribute to the empirical debate over the psychology of moral judgment by standing back from the relevant empirical work, examining the alternative accounts now on the table, and articulating an option that strikes us as in many ways more plausible than the alternatives now being discussed. In engaging in this task, we also call attention to certain data – in this case phenomenological data – that we think ought to figure into evaluating the competing empirical accounts of moral judgment. We think this data provides some *pro tanto* reason in favor of our morphological brand of rationalism.

Here is our plan. In Section 1, we contrast rationalism and intuitionism about moral judgment as these positions seem to be understood by those engaged in empirical research on the topic. In doing so, we focus mainly on the work of Jonathan Haidt, arguably the leader of the recent anti-rationalist, pro-intuitionist trend in moral psychology. Then, in Section 2, we explain how we understand the notion of a moral principle and how we understand the connection between moral principles and reason-giving. In Section 3, we provide a sketch of MR about moral judgment in which we distinguish it from other forms of rationalism, and explain how it can accommodate the data featured in intuitionist views. We next turn to matters of phenomenology in Section 4 – specifically, the phenomenology of intuitive moral judgment-formation, and the phenomenology of moral reason-giving.

Here is where we think morphological rationalism has an advantage over intuitionism, at least of the sort defended by Haidt. We conclude in Section 5.

## 1 Rationalism vs. Intuitionism in Moral Psychology

One of the central questions in the field of moral psychology is: ‘How does moral judgment work?’<sup>1</sup> Rationalists in this field put primary emphasis on moral reasoning, where ‘reasoning’ here is understood as a conscious process that involves steps, “at least two of which are performed consciously” (Haidt 2001, p. 818). Moral reasoning, then, is a conscious process in which one infers a moral judgment from some information that one takes to be morally relevant. Moreover, the claim that this process is *conscious* means that it is “intentional, effortful, controllable, and that the reasoner is aware that it is going on.” (Haidt 2001, p. 818). For the psychological rationalist, then, people typically reach moral judgments about particular cases (henceforth ‘moral judgments’ for short) through a process of conscious reasoning. Moreover, such reasoning presumably involves an appeal to some moral principles or rules in reaching a moral judgment. (For more on this presumption, see Section 2 below.)

Psychological intuitionists, by contrast, emphasize spontaneous gut-level responses—moral intuitions—in answering the central question about how moral judgment works.<sup>2</sup> Haidt and Bjorklund define moral intuition as “*the sudden appearance in consciousness, or at the fringe of consciousness, of an evaluative feeling (like-dislike, good-bad) about the character or actions of a person, without any conscious awareness of having gone through steps of searching, weighing evidence, or inferring a conclusion*” (2007). Haidt’s particular version of psychological intuitionism, the ‘social intuitionist model’ (SIM) features the following core tenets:

- (I) *Intuitionism about moral judgment.* Most moral judgments are the result of the operation of moral intuition.
- (A–R) *Anti-reasoning thesis about moral judgment.* Moral reasoning is not typically part of the moral judgment forming process (at least not in those cases in which moral judgments result from the operation of moral intuition).<sup>3</sup>

<sup>1</sup> The other main question in moral psychology (according to Haidt and Bjorklund 2007) is ‘Where do moral beliefs and motivations come from?’ – a question about how humans come to *acquire* a moral outlook. In response to this question, they discuss empiricist, rationalist, and moral sense theories, and favor a Humean moral sense view. See also Haidt and Joseph (2004).

<sup>2</sup> It is important to note that both rationalist and intuitionist views do (or may) countenance the same sorts of psychological elements that figure into people’s moral psychology. These elements include: various emotions (e.g., disgust, anger), intuitions, moral judgments, and moral reasoning. The difference between these views is how these elements are typically involved in producing moral judgments. According to the rationalist, moral reasoning is at the center of the process, and so it is the proper focus of research in moral psychology. Intuitionism, by contrast, puts primary emphasis on moral intuitions and less emphasis on moral reasoning as a primary focus of research on the production of people’s moral judgments. According to Haidt and Bjorklund, then, the difference between these views is one of emphasis: “Rationalists say the real action is in reasoning; intuitionists say it’s in quick intuitions, gut feelings and moral emotions” (2007).

<sup>3</sup> Haidt does allow an important role in his SIM for moral reasoning (see especially Haidt 2003), but it typically takes place in inter-personal social contexts (hence, the ‘social’ aspect of the model). Haidt again:

The core of the [social intuitionist] model gives moral reasoning a causal role in moral judgment but only when reasoning runs through other people. It is hypothesized that people rarely override their initial intuitive judgments just by reasoning privately to themselves because reasoning is rarely used to question one’s own attitudes or beliefs. (Haidt 2001, p. 819).

- (C) *Confabulation*. The process of moral reason-giving is generally post hoc confabulation: individuals, when they attempt to give reasons for the moral judgments they make, tend to follow a ‘makes-sense’ moral script, offering what they think are the considerations that led them to their judgment, but which actually played no role in producing the moral judgment in question. They confabulate: “moral reasoning does not cause moral judgment; rather, moral reasoning is usually a post hoc construction, generated after a judgment has been reached” (Haidt 2001, p. 814). And so, when faced with a demand for a verbal justification of some moral judgment, “one becomes a lawyer trying to build a case rather than a judge searching for the truth” (Haidt 2001, p. 814).

Thesis A–R is an implication of I together with the definition of reasoning as a conscious, effortful process. However, the connection between I and A–R on the one hand and C on the other seems to be mediated by the following assumption:

- (A) *Assumption*. Unless conscious moral reasoning is part of the process leading to moral judgment, subjects’ reason-giving efforts are not a matter of citing considerations that really did play a causal role in the generation of the judgments in question; rather, reason-giving in these cases is a matter of confabulation.

Haidt does not himself (as far as we know) make this assumption explicit, but it seems clearly operative in his thinking in claiming that reason-giving is most often a matter of post hoc confabulation. And anyway the assumption is initially plausible. After all, if people’s moral judgments are mostly a matter of spontaneous, intuitive response to some actual or hypothetical case, and thus not the result of a process of conscious reasoning, then how could it be that their post-judgment reason-giving activities are really anything other than putting forth considerations, awareness of which they falsely think was involved in the production of their spontaneous moral judgments? How could reason-giving in these cases be other than confabulation?

One possibility is that (1) intuitive moral judgments result from *unconscious* moral reasoning, (2) this happens so quickly that the judgments are experienced consciously as spontaneous even though they really are the product of rapid, multi-step, unconscious deliberation, and (3) the reasoning that operates subconsciously in one’s initial judgment-formation somehow surfaces into consciousness when one subsequently undertakes to give reasons for one’s judgment. This picture of moral judgment-formation seems very implausible, however, in the absence of specific theoretical and/or empirical considerations that might favor it. Unless and until such considerations are put forth, the hypothesis of rapid unconscious reasoning seems little more than an ad hoc way of trying to defend rationalism. We ourselves doubt that suitable theoretical or empirical motivation for this hypothesis will be forthcoming.

We agree with the intuitionist that in a great many cases people’s moral judgments are spontaneous and are not preceded by a process of reasoning (either conscious or unconscious) that produces those judgments. But we are also sympathetic to the rationalist idea that moral principles are involved in the generation of moral judgments. Moreover, we are also skeptical of Haidt’s claim that moral reason-giving is by and large confabulation, though undoubtedly people do sometimes confabulate. So, our plan is to explain why we think A as well as C ought to be rejected. But if one rejects A, what sort of psychological story might

- (1) Embrace what is plausible about intuitionism,
- (2) Avoid the problems with traditional rationalism, and
- (3) Do so without compromising the process of reason-giving?

This question mentions three desiderata for developing a coherent and plausible rationalist alternative regarding the psychology of moral judgment. And these desiderata drive the account we are going to propose.

However, for purposes of clarification, we will first make some preliminary comments about moral principles, noting in particular how our understanding of them relates to moral reasoning in which one gives (to oneself or others) moral reasons. We will also note how our usage is appropriately non-committal with regard to various ways of interpreting the import of moral principles.

## 2 Moral Principles and Moral Reason-giving

Rather than attempt to define the notion of a moral principle (rule), we will simply list some examples and make a few comments about how we understand this notion. Here, then, is a short list of commonly recognized moral principles:

1. Cruelty is wrong.
2. One ought not to kill an innocent human being.
3. One has a duty to avoid lying.
4. Do not commit adultery.
5. One ought to give to charity.
6. Act only on that maxim that you can consistently will as universal law.
7. An action (in some context) is obligatory just in case it would produce a greater amount of net intrinsic value than would any other alternative action open to the agent in that context.

Here we call attention to three points about our understanding of moral principles.

First, note that these sample principles pick out features of an action, at some level of generality (e.g., cruelty, adultery, not consistently willable as universal law), that purport to be morally relevant, counting as reasons for or against certain actions. Generally speaking, moral reasoning (involving reason-giving) about particular cases involves appealing to what one takes to be morally relevant features of the case at hand—features that one supposes provide a deep (enough) reason for the judgment. Such ‘deep’ features are the ones referred to in moral principles. Appealing, then, to such reasons represents an appeal to moral generalizations – to moral principles.

Second, it is common in ethics to distinguish exceptionless principles (however detailed) from *ceteris paribus* principles – the latter often referred to as principles of *prima facie* duty. For our purposes, it does not matter whether an individual accepts a moral principle as holding without exception or as holding only *ceteris paribus*. Our concern is how (if at all) such moral generalizations typically figure into the generation of an individual’s moral judgments.

Third, in recent years, moral philosophers have debated the question of whether there are any features of actions (other than their moral status) that always count as reasons bearing on action (whenever they are instantiated) and always count in the same way – either in favor or against the action in question. Reasons ‘atomists’ claim there are moral reasons that always count and always in the same way; they are opposed by reasons ‘holists.’ We can be agnostic about this dispute for present purposes because it does not affect our question about the *psychological role* of moral principles in generating moral

judgments.<sup>4</sup> Perhaps some individuals take the type of consideration mentioned in our first sample principle, namely cruelty, as always counting against an action whenever the action would be an instance of cruelty. Perhaps some don't, and understand the property of cruelty to be generally, but perhaps not invariantly, of negative moral relevance.<sup>5</sup> Either way, we can ask what role (if any) moral principles play in the production of moral judgments.

In light of these clarificatory remarks, we can refine our main question in the following way: How might one:

- (1) Embrace what is plausible about intuitionism, which denies that moral judgments are reached by a process of moral reasoning employing moral principles, yet
- (2) Maintain the rationalist idea that moral principles are typically involved in the production of moral judgments, and
- (3) All the while maintain that reason-giving need not be (and in general is probably not) confabulation?

Let us proceed to sketch a view that satisfies these three desiderata.

### 3 Morphological Rationalism: A Sketch

Intuitionism in moral psychology is right in holding that a great many moral judgments are spontaneous in the sense of not being preceded by conscious reasoning. But rationalism is also right in holding that people's intuitive moral judgments are often based on reasons – considerations that figure in the generation of these judgments, and do so in a way that turns on a person's acceptance of moral principles. How might these claims be reconciled?

They can't be reconciled unless one rejects assumption A, to the effect that if a process of reasoning is not involved in the production of a moral judgment, then considerations that people are disposed to cite as reasons for their judgments play no role in the causal generation of those judgments. But, then, if one rejects A, how is one to understand the generation of intuitive moral judgments? This is a 'How possible?' question. But, of course, there is also a 'How plausible?' question to be answered: even if one can sketch a possible position that rejects assumption A, there is still the question: How plausible is this sort of position vis-à-vis its competitors? In what immediately follows, we take up each question in order. Then, in the following section, we continue our case for the plausibility of our view by turning to matters of moral phenomenology.

<sup>4</sup> So, it is important to notice that the present issue regarding the roles of principles in moral psychology is distinct from (and largely orthogonal to) philosophical issues concerning the metaphysics and epistemology regarding moral principles. Thus, even if one is a metaphysical holist about moral reasons and so, along with Jonathan Dancy (2004) and David McNaughton (1988), denies that there are any considerations that are invariant in their moral relevance, and even if one is an epistemological particularist (Garfield 2000) and denies that moral principles are first in the order of moral justification and knowledge, one can still allow that moral principles (understood as generalizations) might still play the sort of guiding role via proceduralization that we explain in the next section.

<sup>5</sup> It is worth noting that it is possible (and we think plausible) to accept the idea that moral principles do contain a 'ceteris paribus' clause, even though the morally relevant features mentioned in such principles can be silenced and reversed, and hence need not be invariant in their relevance or valence.

### 3.1 How Possible?

According to morphological rationalism, moral principles typically guide without being represented by the agent. We say that morally mature individuals typically possess such principles (their contents) *morphologically* and that the manner in which such principles are operative in producing particular moral judgments is by being *proceduralized*. So to sketch our view, let us introduce and illustrate these two main ideas.

Moral principles have intentional content. In particular, they specify considerations that count in favor of or against courses of action. They also can and do function to guide one's moral judgments about what one ought or ought not to do. Here, we are particularly interested in their judgment-guiding role. So, how might one understand the idea *possessing a moral principle in a manner that is operative in generating moral judgments*? There are three main ways of possessing content: one may possess it occurrently, or in a dispositionally indirect manner, or in a dispositionally direct manner. To possess content in this third way is to possess it morphologically. Let us consider each such manner of possession in order.

For an individual to possess a moral principle (its content) *occurrently* at a time is for a token of the content state type to occur as a state or event within the person at that time. If one possesses a moral principle in this manner, then it is obviously ready for use in arriving at content-relevant moral judgments at that time.<sup>6</sup> My occurrently believing on some occasion that Jones lied to Smith in order to swindle her out of her inheritance, together with my occurrently representing a moral principle (that I accept) to the effect that lying for gain is morally wrong, results in my judging that Jones acted wrongly.

For an individual to possess a moral principle (its content) in a dispositionally *indirect* manner is for the individual to be disposed to produce a token state or event that is an explicit representation of that principle, which then plays a role in generating a moral judgment in accordance with the principle. (This might happen either consciously or unconsciously.) So what makes this manner of dispositional possession *indirect* is that the link between the disposition and the moral judgment is *mediated* by an occurrent representation of the principle. We can repeat the Jones lying to Smith example to illustrate how this manner of possession can be operative in the generation of token moral judgments on some occasion. Just as in the case of occurrent possession, indirect dispositional possession requires that in order to be operative on some occasion, a moral principle (accepted by the individual) must be *occurrently represented* by the individual on that occasion.

Call any view according to which moral judgments are typically the causal result of the operation of principles that the individual occurrently represents, *representational rationalism*. Morphological rationalism contrasts with the representational variety.

For an individual to possess a moral principle (its content) *morphologically* is for the individual to be disposed to undergo transitions in cognition, from certain input to moral judgments as output, such that (1) these transitions systematically and non-accidentally conform with the moral principle, (2) this systematic conformity results from the person's persisting psychological structure (the morphology of the cognitive system), and (3) those cognitive transitions typically result from this persisting structure *without the mediation of a tokened representation of the principle*. In this case, the manner of possession is

<sup>6</sup> At any rate, it is obviously ready for use if it is occurrently tokened in a suitable way. Insofar as one construes human cognition on the model of a computer (a Von Neumann machine), one might think of such suitable tokening as occurring within the central processing unit, and not just in the cognitive system's memory banks.

dispositionally *direct*. One may think of this sort of possession as a matter of *know how* – a skill that is or has become part of the individual’s repertoire for negotiating her social world. When a principle or norm is possessed morphologically, one can say that its manner of operation is *procedural* – in virtue of possessing the principle in this manner, an individual is disposed to form moral judgments that non-accidentally conform to the principle.<sup>7</sup> Moral judgments are thus formed ‘automatically’ and spontaneously in virtue of the individual’s persisting psychological structure as a morally competent individual. One may call such automatic know-how ‘procedural knowledge.’

In the case of Jones’ lie, in coming to believe that Jones told the self-benefiting lie in question, one comes to believe that what Jones did was wrong. But one need not, and typically does not, occurrently represent a content-relevant moral principle that connects actions under a certain non-moral description with the concept of moral wrongness. Similarly, in hitting a tee shot, Tiger Woods need not occurrently represent any principles of such golf shots that he knows (head down, even balance between both feet, etc.) and that he has thoroughly internalized. Rather, these principles are part of Tiger’s know-how as a golfer and operate procedurally. So, both Jones and Woods have procedural knowledge of various principles that operate automatically to bring about relevant judgments (in the case of Jones) and behavior (in the case of Woods).

Both dispositionally indirect and morphological ways of possessing a moral principle involve dispositions. The difference is that the former involves an indirect disposition to generate a corresponding moral judgment – a disposition to generate a mental representation of the principle which then plays a role in generating moral judgments with appropriate content. Morphological possession of a moral principle, by contrast, involves a direct disposition to generate a corresponding moral judgment, without the exercise of that disposition being mediated (either consciously or unconsciously) by an occurrent representation of the principle in question. Still, with this morphological manner of possession, moral principles are properly said to be operative in generating specific moral judgments. One comes to have a particular moral judgment on some occasion *because* one accepts the relevant moral principle. One has internalized the principle morphologically, and one judges as one does *because* one has thus internalized the principle.<sup>8</sup>

<sup>7</sup> As we explain in the next section, there are other dispositions characteristic of possessing a moral principle morphologically.

<sup>8</sup> Are such explanations of moral judgments in terms of morphologically possessed moral principles in danger of being trivial? One way in which this triviality worry might be pressed is by noting that in every case in which someone comes to have or make a moral judgment, it is trivially true that there is a moral principle that can be invoked to help explain the particular judgment. After all, in coming to hold a moral judgment about a particular case in which the judge is responding to some particular set of perhaps very complex circumstances, it will be possible in principle to formulate a *very* specific moral ‘generalization’ that specifies in its antecedent as much information as you please about the factual details of the case (the circumstances and action) and specifies in its consequent the moral feature or property being attributed by the judge to the specific act token described in the antecedent. So, according to this conception of a moral principle, it is completely trivial that moral principles are part of the explanation of a person’s particular moral judgments. And so it would be if one conceives of moral principles in this way. We don’t. Rather, we have in mind explanatorily interesting moral generalizations that are suitable for helping to explain a robust range of actual and counterfactual judgments a person is disposed to make. For instance, there is considerable empirical evidence that people attribute moral responsibility to agents whose actions produce harm to someone based on a principle that takes into consideration such factors as intention, negligence, recklessness, and foresight. A person is not held morally responsible if the harm caused was accidental (without intention, negligence, recklessness) or done involuntarily (e.g., under duress) or without foresight of the resulting harm. (On this matter, see Darley and Schultz 1990.) And it is a genuine non-trivial matter, according to this conception of moral principles, to appeal to such generalizations, whether consciously or unconsciously operative, in explaining people’s moral judgments.

We can now formulate the main claim of morphological (moral) rationalism:

- (MR) *Morphological Rationalism*. Moral principles may be, and typically are, possessed by morally mature individuals morphologically, and the morphological possession of moral principles then *procedurally guides* the ‘automatic’ formation of particular judgments. Such principles, to be procedurally operative on some occasion, need not be explicitly represented by some occurrent mental state in the individual on the occasion in question.

Notice that MR is compatible with the intuitionist and the anti-reasoning theses (I and A–R respectively) featured in Haidt’s SIM. That is, MR is fully compatible with the idea that particular moral judgments are typically intuitive – not preceded by conscious deliberation – and thus not preceded by a process of conscious reasoning.

With our version of rationalism on the table, let us consider what can be said for it.

### 3.2 How Plausible?

We just mentioned that MR fits comfortably with two of the theses featured in Haidt’s SIM, and both of those theses are well-supported by psychological research on automaticity. (See Bargh and Chartrand 1999). Moreover, the automatic operation of principles in other norm-guided areas of thought has gained strong support. For instance, in epistemology, John Pollock and Joseph Cruz have argued that the central concept of epistemic justification is normative – it concerns the epistemic permissibility of beliefs – and that epistemic norms describe when it is permissible to hold a belief. Further, on their view, a justified belief is one that is guided by and thereby complies with epistemic norms. But how do epistemic norms perform their guiding function? Pollock and Cruz reject what they call the ‘intellectualist’ model according to which norms operate by being explicitly represented by the cognizer and then applied to the task at hand. Instead, they defend the idea that such norms are ‘procedural’ in the way they operate. Here is how they summarize this idea:

The internalization of norms results in our having “automatic” procedural knowledge that enables us to do something without having to think about how to do it. It is the process that we are calling “being guided by the norm without having to think about the norm.” This may be a slightly misleading way of talking, because it suggests that somewhere in our heads there is a mental representation of the norm and that mental representation is doing the guiding. Perhaps it would be less misleading to say that our behavior is being guided by our procedural knowledge and the way in which it is being guided is described by the norm. What is important is that this is a particular way of being guided (Pollock and Cruz 1999, p. 128).

Notice that the authors make a point of saying that norm guidance need not proceed via a token mental representation of the norm for it to be operative in the generation of justified beliefs. And, of course, this is the main idea involved in possessing a moral principle (its content) morphologically.

In developing their account of epistemic norms as procedural, Pollock and Cruz make use of a widely accepted account of grammatical competence championed by Chomsky. According to Chomsky, people’s judgments of grammaticality are to be explained by their possession of grammatical rules – rules which individuals may not be able to explicitly formulate, but to which their competent grammatical judgments conform. The rules in

question are plausibly understood as ‘procedural,’ know-how rules which operate automatically – without needing to be represented by the competent users of the language.<sup>9</sup>

So, the plausibility of these accounts regarding the operation of epistemic norms and grammatical rules is inherited by our account of the operation of moral principles. Of course, we haven’t addressed matters of detail in our sketch. Of particular note are details concerning how one comes to possess moral principles, and how the operation of such principles in generating intuitive moral judgments comports with the fact that such judgments are often accompanied by affect-laden emotions. But our view leaves such matters of detail completely open. For instance, according to innatist views of morality, such principles are part of our genetic endowment. And if so, then it is plausible that they operate procedurally. By contrast, according to empiricist views, moral principles are learned. But again, learning can result in principles coming to be proceduralized. Mixed views of the sort found in Haidt and Joseph (2004) combine elements of innatism and empiricism. MR is compatible with all three views.

A similar remark applies to issues about emotions and intuitive moral judgments. One would expect the automatic operation of moral principles to be associated with certain emotional responses. And this thought can be accommodated by morphological rationalism in a variety of ways. For instance, on a ‘sentimental rules’ account of moral judgment of the sort defended by Shaun Nichols (2004), moral judgments are the partial product of the operation of moral rules (principles) that are backed by sentiments. The sentimental backing is supposed to explain various features of moral experience, including the differences in seriousness, applicability, authoritativeness, and types of backing reasons between how one experiences moral violations as compared to violations of mere conventions. Marc Hauser (2006) has developed an account of the psychology of moral judgment that makes explicit use of Chomskian linguistic theory. On Hauser’s view, one comes equipped with a ‘moral grammar’ composed of a set of moral principles that guides our moral judgments in roughly the same automatic way in which innate rules of general grammar guides the production of judgments about grammatical sentences. On Hauser’s particular view, affect-laden emotions are typically triggered by a moral judgment, rather than being involved in their generation. For both Nichols and Hauser, then, both principles and emotions are involved in coming to have or make moral judgments, though their views differ in the details of how such judgments are produced. Morphological rationalism is compatible with both of these views, since they both involve moral principles as part of the genesis of typical moral judgments.

Having sketched our view and pointed to some of the psychological issues it leaves open, let us pause to take stock. Earlier we announced three desiderata we think a plausible account of the psychology of moral judgment ought to satisfy. Such an account ought to (1) embrace what is plausible about intuitionism; (2) avoid problems with traditional rationalism; and (3) do so without compromising the integrity of moral reason-giving. Our morphological rationalism, as so far presented, satisfies the first two. MR is fully

<sup>9</sup> Admittedly, one possibility is that the rules of grammar are explicitly represented *unconsciously* during language processing, and that the exercise of grammatical competence during speech production and speech recognition is a matter of unconscious, extremely rapid, application of these explicitly represented rules. But surely a more efficient and more effective way for language processing to work would be via the automatic guidance of proceduralized grammatical rules – say, via innate proceduralized rules of universal grammar together with learned proceduralized rules of one’s specific language. For a book-length defense of the claim that *much* of human cognition needs to operate procedurally (on pain of encountering computational intractability otherwise), see Horgan and Tienson (1996). This book includes a proposal for understanding procedurally-operative morphological content by appeal to the mathematics that goes naturally with connectionist models, viz., dynamical-systems theory.

compatible with the idea that such judgments are very often psychologically spontaneous; it does not require that competent moral judgments are the result of a process of reasoning. Moreover, MR fits nicely with how principles seem to guide other realms of thought involving one's know-how. Let us now turn to the third desideratum concerning the practice of reason-giving.

#### 4 The Phenomenology of Moral Judgment and Moral Reason-giving

Recall the dialectical state of play that is propelling this discussion. Haidt accepts the intuitionist thesis I (people's moral judgments are generally the result of spontaneous intuitions) and he accepts the associated anti-reasoning thesis, A–R, that reasoning does not play a role in generating such judgments. Together with assumption A (according to which (roughly), if people's moral judgments are not the product of reasoning, then their practice of reason-giving is post hoc confabulation) he is led to the confabulation thesis, C: the considerations people sincerely bring forth as reasons for their antecedent moral judgments are not playing a causal role in generating those judgments; rather, in giving moral reasons, people are guilty of confabulation. No doubt people do confabulate. But for Haidt, confabulation in the practice of moral reason-giving is the rule, not the exception. This strikes us as implausible in the extreme.

MR is not driven to embrace Haidt's confabulation thesis because our view allows for, and indeed makes sense of the rejection of A. This means that our view does not require the counter-intuitive claim that moral-reason-giving is generally confabulation. Part of the counter-intuitiveness of the confabulation thesis is that it conflicts with certain facts about people's moral phenomenology. We claim that facts about the phenomenology of moral judgment-formation and moral reason-giving constitute *data* that ought to constrain theory acceptance in moral psychology: an adequate moral psychology should *explain* such data. Furthermore, we claim that a methodological default assumption concerning such phenomenological data – especially when the pertinent phenomenology is common across individuals, and is robustly persistent – is that the right psychological explanation should appeal to facts about human moral competence. I.e., explanatory practice in psychology should conform to the following *maxim of default competence-based explanation* (for short, the DCBE maxim):

(DCBE) All else equal, a theoretical explanation of a pervasive, population-wide, psychological phenomenon will be more adequate to the extent that (1) it explains the phenomenon as the product of cognitive competence rather than as a performance error, and (2) it avoids ascribing some deep-seated, population-wide, error-tendency to the cognitive architecture that subserves competence itself (e.g., an architecturally grounded tendency to erroneously conflate post-hoc confabulation with articulation of the actual reasons behind one's moral judgments).<sup>10</sup>

<sup>10</sup> Certain specific kinds of error or error-tendency might be an inherent product of a cognitive system's competence, rather than resulting from breakdowns in that competence. (For example, perhaps an ideally competent human visual-perception system will be subject to the Muller-Lyer illusion, by virtue of the system's innate cognitive architecture.) Clause 2 makes explicit that theoretical hypotheses about cognitive architecture are themselves subject to the DCBE maxim – i.e., all else equal, a hypothesis about cognitive architecture is better insofar as it does *not* attribute some systematic and pervasive form of error-tendency to the architecture itself.

Given this sensible-looking methodological principle, moral phenomenology favors non-confabulation; therefore, at least on this score, our view is preferable to Haidt's intuitionism. But because our view also accommodates what is plausible about Haidt's, our version of morphological rationalism is overall preferable to his. Let us proceed to consider the matter of moral phenomenology more closely, while holding in view the DCBE maxim.

To make our case, we begin by calling attention to the sorts of dispositions involved in the morphological possession of a moral principle. In the previous section, we stressed the direct disposition to spontaneously form content-relevant moral judgments. But, this is only a partial characterization; there are other direct dispositions that one possesses in virtue of possessing a moral principle morphologically. Here, then, is a more robust dispositional characterization of what is involved in possessing a moral principle morphologically.

Morphological possession of a moral principle involves, *inter alia*:

- A direct disposition to form content-relevant moral judgments,
- A direct disposition to experience one's moral judgment as fitting the circumstances to which one is responding in making or having the moral judgment,
- A direct disposition to present reasons for one's moral judgment if one is prompted to do so,
- A direct disposition to experience one's reason-giving as fitting smoothly with the moral judgment for which one is giving reasons.

The direct dispositions to experience the fittingness of one's moral judgments to one's circumstances, and to experience one's reason-giving as fitting the corresponding moral judgments, concern those aspects of moral phenomenology to which we wish to call special attention. In particular, we call attention to the *non-jarring* character of such experiences. Let us consider, in turn, first the non-jarringness of intuitive moral judgments themselves, and then the non-jarringness of reason-giving experiences – in both cases, bearing in mind the DCBE maxim.

### 1. The Non-Jarringness of Intuitive Moral Judgments

One's on-the-spot intuitive moral experiences typically involve a sense (even if not the focal point of one's attention) of responding in a fitting, non-jarring, manner to what one takes to be good reasons. In seeing an elderly man slip and fall on an icy sidewalk, one experiences the morally salient aspects of the circumstances as normatively 'calling for,' and hence being a good reason for, a certain response that one takes to be fitting – in this case rendering aid. And more particularly, one experiences oneself as responding to what is in this specific circumstance *these particular here and now* reasons – as responding to *that-fallen-man-right-over-there-needing-my-help*. One's experience of the circumstances calling for a certain response fit together with one's intuitive judgment that one ought to help and one's swinging into action. Contrast a bizarre case in which one finds oneself responding to the fallen man by having the thought (out of the blue) that one ought to teach the gentleman Spanish. Such a thought (without any appropriate prompting) would strike one as alien, a case of what is called 'thought insertion.'<sup>11</sup>

How might one best explain the phenomenology of non-jarringness that typically accompanies intuitive moral judgments – the phenomenology of being reason-based and situationally appropriate, as opposed to the phenomenology of being "bizarrely out of the

<sup>11</sup> On the phenomenon of thought insertion, see Stephens and Graham 2000.

blue”? Here the DCBE maxim comes to the fore. All else equal, an explanation that accommodates this population-wide, robust, phenomenology by construing it as veridical – i.e., that treats the phenomenology as *competence-based* – will be preferable to an explanation that construes the phenomenology as falsidical – i.e., that treats it as a robust, population-wide, *performance error* (or as a robust, population-wide, error-tendency in cognitive architecture itself). Morphological rationalism offers the methodologically preferable form of explanation, whereas Haidt’s theoretical stance requires the claim that the phenomenology of reason-based appropriateness is a systematic illusion. Morphological rationalism thus yields the *better* explanation, unless and until the advocates of the competing explanatory hypothesis can cite specific empirical data that is sufficiently powerful to warrant an appeal to an explanatory hypothesis that is error-based rather than competence-based. (In principle, empirical data could be provided that would over-rule the DCBE maxim and would favor an error-based explanation; but the burden of proof on the error theorist is very heavy indeed – especially when the alleged error is claimed to be both population-wide and robustly persistent.)

## 2. The Non-Jarringness of Reason-Giving Experience

The general point about the non-alien, non-jarring character of moral experience is also reflected in certain phenomenological facts about the activity of reason-giving. In providing reasons for a past judgment, whether to oneself or others, there is a contrast between cases in which one experiences one’s reasons as non-jarring and cases in which one experiences them as jarring – not quite fitting the judgment for which one is offering reasons. So let us consider in more detail the phenomenology of reason-giving. Here we call attention to a number of points. (Our discussion of the non-jarringness of reason-giving experience will be more extensive than what we said about the non-jarringness of intuitive moral judgments themselves – in part because the phenomenology of reason-giving is more directly relevant to Haidt’s contention that people typically confabulate when seeking to articulate a rationale for their moral judgments.)

First, as explained in Section 2, the practice of moral reason-giving is typically a matter of citing morally relevant considerations that one generally takes to be defeasibly sufficient for an action’s all-in moral status – considerations featured in moral principles.

Second, in sincerely giving reasons for one’s all-in moral judgments, one typically experiences the giving of reasons as fitting smoothly with the experiences in which those judgments were formed, and as helping to make sense of those experiences. The contrast is with cases in which the sincere citing of what one takes to be reasons strike one as jarring – as not really fitting with the original experience. Consider, for example, the experience of sipping a wine during a conversation and thinking it to be refreshing. When asked why, one might say that it is has a slight effervescence, but find one’s saying so somehow not quite right – not quite getting at what one’s judgment was responsive to in tasting the wine. Another sip confirms the slight sense of non-fit or jarringness between one’s spontaneous judgment and one’s proposed reason for that judgment. Attending more carefully to the wine’s qualities, one corrects one’s initial reason by thinking that it is the wine’s crispness that best fits with and explains one’s judgment that the wine is refreshing. In fact, upon taking a second sip, one comes to think that the wine isn’t really effervescent. Now consider a non-moral case that is perhaps closer to moral judgment. Upon viewing a statue of a Greek god, one is immediately struck by its beauty, and upon reflection one thinks that it’s beauty is due primarily to the depiction of the proportion of the various body parts – tacitly appealing to the idea that proportion defeasibly underlies the beauty of a this sort of statue. As one continues to scrutinize the statue, one’s reason-giving hypothesis does not quite feel right in the sense that it doesn’t seem to pinpoint what is driving

one's spontaneous aesthetic judgment. Upon further reflection, one hypothesizes that it is the statue's grace that has triggered one's original judgment. In citing this reason as a basis for one's judgment, there is a feeling of satisfaction: one's reason-giving phenomenology fits with one's ongoing experience, including one's judgment of the statue's beauty.

This same phenomenology is present in connection with moral reason giving. Here are three illustrations.

#### 4.1 Harman's Hoodlums

Recall Gilbert Harman's case of rounding a corner and seeing some hoodlums pour gasoline on a live cat and ignite it. As Harman observes: "you do not need to *conclude* that what they are doing is wrong; you do not need to figure anything out; you can *see* that it is wrong" (Harman 1977, p.4). In explaining one's reasons for this particular moral judgment about the wrongness of this action, one cites the morally salient consideration featured in the principle proscribing cruelty. And perhaps one will also cite other reasons. One's citing such reasons for this judgment is experienced as getting at and hence as fitting the original judgment.

#### 4.2 Connie's Conversation

Connie and Audrey have enjoyed a close friendship spanning three decades. Despite living in different parts of the world for the past 12 years, their friendship has not eroded. Email has made staying in touch on a weekly, sometimes daily, basis easy even given their family lives and heavy work schedules. Connie has just received exciting news that she is to be promoted from a director's position to a vice president of the corporation for which she has worked for nearly 25 years. This is the sort of news befitting a phone call to Audrey. She calls. But instead of being greeted with Audrey's typical hearty and cheerful "Hello? Audrey speaking," she hears a weak and decidedly downbeat, "Hello." Sophie hesitates a moment: "Hello?, Audrey is that you? This is Con." ...Short silence... "Are you there? Are you alright?" Audrey: "Yes, it's me, Connie, but I'm afraid I have some bad news...I've just been let go...my job, I mean." Connie: "My God! I'm so sorry! This is terrible news!"

This begins a long conversation between Connie and Audrey during which Connie is mostly silent allowing Audrey to express her disappointment, embarrassment, and anger over her job loss. Early on in this conversation it is patently clear to Sophie that now is obviously not the time to share her good news with Audrey – she should wait for a later, more appropriate occasion. Her judgment about waiting is psychologically immediate: she doesn't consciously rehearse or weigh various considerations; rather, once she hears the voice of Audrey and what Audrey has to say, her judgment to withhold the news is spontaneous. After hanging up, Connie recounts how she felt during the conversation as well as *why* she was right not to speak up about the job: it would have been callous of her to mention it, and particularly so given that she was speaking to Audrey.

#### 4.3 Ian's Intervention<sup>12</sup>

Ian is hiking alone through mountainous terrain, lost in thought, when he hears the not so distant voices of a man and a woman quarreling. Ian climbs to the top of a large rock where

<sup>12</sup> This example is inspired by a scene from Ian McEwan's novel *Amsterdam* (1999, ch. iii) in which composer Clive Linley is confronted with the sort of choice we go on to describe. In our story, Ian does the right thing; the incident in McEwan's story might be called 'Clive's callousness.'

he can now see the quarreling couple standing face to face in a small clearing about 50 yards away. As the confrontation continues, loud talking soon gives way to shouting. The man grabs the woman's elbow pulling her in his direction. She breaks loose with a sharp downward jerk of her arm, and turns to run, but the man tackles her from behind. They fall to the ground; the woman trying to crawl away, the man holding onto her ankles. Ian imagines running to help the woman, thinking that upon reaching them, the man might run off, or the man might turn on him. The man, having gotten up is now dragging the woman both hands on her left ankle; she screams. Ian knows what he must do. He jumps down from the rock and runs to help, hoping his intervention will cause the man to flee.

Once he understands that the man's aggression toward the woman is serious, Ian's judgment about what he must do is psychologically immediate. After the incident Ian reflects on his judgment and consequent behavior thinking that he had to intervene: the woman was in danger, he was the only other person around (so far as he knew) and it would have been cowardly of him not to intervene. He experiences his reason-giving as fitting with and explaining his moral ought-judgment.

In all three cases, we have individuals who come to have or make an intuitive moral judgment for which they provide (to themselves) reasons for their judgment. The activity of reason-giving (or reason-rehearsing) is experienced as fitting smoothly–non-jarringly – with their overall moral experience including, in particular, the moral judgments they make on that occasion. Moreover, were they prompted by someone (soon after these experiences) to explain why they made the judgment they did, these individuals would presumably be able to effortlessly articulate those considerations that they take to be their reasons – reasons that not only justify the judgment but also figure into an explanation of why they made the judgment they did.

Contrast these cases with the phenomenon that Haidt (2001) calls 'moral dumbfounding.' His prime example is subjects' responses to an incest scenario between brother and sister who take precautions against pregnancy, enjoy the experience, and who decide not only to keep their sexual encounter a secret, but decide to let their encounter be a one time thing. In this case, subjects who make spontaneous moral judgments about this hypothetical case find themselves embarrassed and flustered in offering reasons for the judgment in question – reasons that upon little reflection they realize are contradicted by the facts of the case. As Haidt reports, in such cases subjects experience their reason-giving activity as jarring, not fitting with the judgment. The 'reasons' they initially offer for their moral judgment about the case – the judgment that what the siblings did was wrong – are confabulations.

Recall now the DCBE maxim. We don't see any reason to think that the subjects in our three cases of moral reason-giving are confabulating. Certainly, they would not be dumbfounded were they to be asked about their reasons for the judgments they make. And the reasons they give would likely survive their own critical scrutiny. Of course, even so, they *could* be confabulating; it might be that what they cite as reasons are not really causally operative in generating their moral judgments. But why think this about these and many other such cases? One would be *forced* to suppose that the subjects must be confabulating were one to accept assumption A. But we have already explained in the previous section why this assumption is particularly dubious.

The non-jarring character of many moral reason-giving experiences is psychological data that should be taken into account by any theoretical account of the production of moral judgments. Indeed, this phenomenology of non-jarringness is data that needs explaining. And, so we claim, it is better if this data can be taken at face value – as the product of subjects' cognitive *competence* concerning matters of moral judgment, rather than as the

product of a (putative) robust, population-wide, tendency toward confabulation. This brings us back to our view and its plausibility vis-à-vis competing views.

As mentioned in Section 1, assumption A seems to flow from the perhaps natural thought that if people's intuitive moral judgments don't result from a process of reasoning in which a principle is represented and applied to the case being judged, how could it be that people's post-judgment activities of reason-giving are anything but confabulation? MR provides the basis for an answer to this question. In the previous section, we stressed one characteristic sort of disposition involved in the morphological possession of a moral principle: an individual's being non-accidentally disposed to judge in accord with that principle. But, as lately mentioned, there are other dispositions involved in such possession of moral principles that are reflected in the phenomenological observations we have just been describing. First, there is the disposition to treat as reasons the sorts of considerations that the moral principles treat as reasons in cases where the moral principle is operative in generating moral judgments. Second, there is the disposition to experience the moral judgment as phenomenologically non-jarring in the circumstances (recall the falling man case). And third, there is the disposition to experience one's activity of reason-giving as fittingly and thus non-jarringly related to the experience of one's having or making the corresponding moral judgment.

Noting and illustrating these dispositions characteristic of the morphological possession of a moral principle (in addition to the disposition of the individual to make spontaneous judgments about particular cases that conform to the principle in question) provides a richer account of what it is to possess a principle morphologically.

## 5 Conclusion

This completes our *prima facie* case in favor of morphological rationalism. MR, unlike representationalist versions of rationalism, does not require individuals to represent principles in order for them to be operative in generating moral judgments. MR also fully accommodates the intuitive, spontaneous nature of many instances of coming to have a moral judgment. And finally, MR makes sense of the non-jarring nature of both (1) spontaneous moral-judgment experiences, and (2) experiences of post-judgment reason-giving.

MR is an empirical thesis about the generation of moral judgments. Various sorts of experimental empirical data will be relevant to evaluating MR and its competitors. But, among the empirical data that bear on MR are the data of first-person moral experience – one's phenomenology of moral experience, including the phenomenology of moral reason-giving. As far as the phenomenological data goes, MR is more plausible than competing views (because it accommodates such data in a way that conforms with the DCBE maxim, whereas views that appeal to massive error do not), and is certainly worthy of consideration as an account of how moral judgment works. In addition, MR is plausible on independent grounds (apart from phenomenology): procedural guidance by morphologically internalized rules is just a lot more resource-efficient, and a lot more rapid, than processing that employs explicit representations of those rules and explicit applications of the rules to concrete cases.<sup>13</sup>

<sup>13</sup> An earlier version of this paper was presented to the 2006 British Society for Ethical Theory meeting in Southampton, England. We thank the audience at this meeting for very helpful comments. We also thank our colleagues Michael Gill and Shaun Nichols for their valuable advice and comments on this project.

## References

- Bargh JA, Chartrand TA (1999) The unbearable automaticity of being. *Am Psychol* 54:462–479
- Dancy J (2004) *Ethics without principles*. Oxford University Press, Oxford
- Darley JM, Schultz TR (1990) Moral rules: their content and acquisition. *Annu Rev Psychol* 41:525–556
- Garfield J (2000) Particularity and principle: the structure of moral knowledge. In: Hooker B, Little M (eds) *Moral particularism*. Oxford University Press, Oxford and New York, pp 178–204
- Haidt J (2001) The emotional dog and its rational tail: a social intuitionist approach to moral judgment. *Psychol Rev* 108:814–834
- Haidt J (2003) The emotional dog gets mistaken for a possum. *Rev Gen Psychol* 8, 283–290
- Haidt J, Bjorklund F (2007) Social intuitionists answer six questions about moral psychology. In: Sinnott-Armstrong W (ed) *Moral psychology, vol 3. The neuroscience of morality: emotion, disease, and development*. MIT Press, Cambridge, Mass
- Haidt J, Joseph C (2004) Intuitive ethics: how innately prepared intuitions generate culturally variable virtues. *Daedalus* p 55–66
- Harman G (1977) *The nature of morality*. Oxford University Press, Oxford
- Hauser MD (2006) *Moral minds: how nature designed our universal sense of right and wrong*. Harper Collins, New York
- Horgan T, Tienson J (1996) *Connectionism and the philosophy of psychology*. MIT Press, Cambridge, Mass
- Kohlberg L (1981) *Essays on moral development, vol 1*. Harper & Row, San Francisco
- McEwan I (1999) Amsterdam. Doubleday, New York
- McNaughton D (1988) *Moral vision*. Blackwell, Oxford
- Nichols S (2004) *Sentimental rules*. Oxford University Press, Oxford and New York
- Piaget J (1932) *The moral judgment of the child*. Free, New York
- Pollock J, Cruz J (1999) *Contemporary theories of knowledge, 2nd edn*. Rowman & Littlefield, Lanham, MD
- Stephens GL, Graham G (2000) *When self-consciousness breaks: alien voices and inserted thoughts*. MIT Press, Cambridge, Mass