

The Moral Cognition/Consciousness Connection

Mark Phelan · Adam Waytz

Published online: 15 September 2012
© Springer Science+Business Media B.V. 2012

People experience joy and pain. We appreciate brilliant colors. We have awareness of our own mental states, and we attribute mental states to others—not only beliefs and desires, but also phenomenal experiences (such as joy and pain). We knowingly construct conceptions of ourselves and deploy these self-concepts to navigate the world around us. These are fundamental aspects of the human experience, encompassing phenomenal consciousness, on the one hand, and self-consciousness on the other. Let us refer to these collectively as consciousness.

Just as we cannot but enjoy conscious experiences, we also naturally conceive of the world in ethical terms with an understanding of right and wrong. Moral violations are continually readily apprehensible; but when it is not obvious that some person should or should not have acted in some particular way, just how they should have acted is often a matter of intense debate. Most of us concede that it would have been wrong for George to shoot Trayvon without provocation, but not if George was violently assaulted after having just stepped out of his car to ask a few questions. But why was George trailing Trayvon for so long? Why was he so intent in knowing what this person was doing in his neighborhood? Doesn't that display a morally suspect character? Such sensationalized news stories are the fodder of conversations over dinner or on the social network, but in ordinary contexts, often when alone, we cannot but assess the actions of others according to a moral framework. We conclude that it was wrong for that patron to have been so curt with the waiter. We assess that it was good for the young woman to have given up her seat for the octogenarian. These assessments apply not only to our fellows' actions, but bleed over onto the actors themselves. That the driver cut off the cyclist on a rainy day reveals an inconsiderate villain who cannot easily occupy the mindset of his compatriots. The cyclist is the clear victim here. We feel her pain.

M. Phelan (✉)
Lawrence University, Appleton, WI, USA
e-mail: mark.phelan@lawrence.edu

A. Waytz
Northwestern University, Evanston, IL, USA

How are these two fundamental aspects of human experience—consciousness and moral cognition—related to one another? At first one might have thought not at all. After all we can clearly make moral assessments when we are calm and collected, beyond the momentary flash of rage or appreciative delight. Indeed, assessments made under such cool conditions may in some sense be ideal. Likewise, our appreciation of moral victims seems to outstrip our attribution of phenomenal states. Pollution, for example, seems wrong first and foremost because it harms a non-sentient environment. Finally, the apprehension of moral situations far removed from us seems to suggest that not even self-consciousness is essential to moral cognition.

But there are also reasons for positing a connection between consciousness and moral cognition. Many traditional philosophical accounts tie interests, rights, and other moral properties to the capacity for phenomenal consciousness. Singer (1975), for instance, argues that, “the capacity for suffering and enjoying things is a prerequisite for having interests at all, a condition that must be satisfied before we can speak of interests in any meaningful way.” And this connection between interests and phenomenal consciousness proliferates. Instrumentalists, for example, ground the value of rights in their role “as instruments for achieving an optimal distribution of interests” (Wenar 2010). At the same time, Kantian moral philosophy ties moral duty to a rich conception of the self as rational agent (Kant 1785/1998). Likewise, psychological research has uncovered important connections between moral cognition and phenomenal state attribution. Blair et al. (1995), for example, find differences in the way ordinary people and psychopaths (with their notorious deficits in moral cognition) attribute guilt. Others have demonstrated that our explanation of our own actions vary depending on whether those actions were good or bad (Miller and Ross 1975). Thus, our self-conceptions are presumably shaped in part by moral cognition, and our self-consciousness informs our moral assessments.

Such considerations provide reasons for a *prima facie* supposition of some connection between consciousness and moral cognition. But what is the nature of that connection? And how tightly bound are these two essential aspects of the human experience? This special issue of *Review of Philosophy and Psychology* explores these questions. In this introduction, we will first clarify what exactly we mean by consciousness (a notoriously multifaceted term). We will then briefly introduce moral cognition and investigate some recent accounts of ways in which consciousness (in its many forms) and moral cognition intersect. Against the backdrop of these reflections, we will ultimately introduce the topics discussed in this issue, which push research at the juncture of consciousness and moral cognition in exciting new directions.

1 Consciousness

It is obviously not possible to satisfactorily survey the long history of consciousness research in an introduction of this length. However, to appreciate the contribution of many of the papers in this issue, it is essential to distinguish between several different mental phenomena, which fall under the common heading “consciousness”. Many philosophers have surveyed the different concepts that share this label. The discussion

below draws primarily from David Rosenthal (1986) and Ned Block (1995) (see also Van Gulick 2011, for a more thoroughgoing taxonomy):

Responsiveness:

One basic use of the term consciousness corresponds to awakesness, or responsiveness to the external world. A patient emerging from sedation is said to be conscious in this sense. What's more, this state seems to apply to all manner of organisms. It is a matter of debate whether lobsters feel pain, and they are quite generally denied self-awareness, nonetheless a sedated lobster could be said to be returning to consciousness in this sense. What responsiveness seems to presuppose is that the entity is aware of its environment. It thus seems to involve a variety of transitive consciousness.

Transitive Consciousness:

In addition to regarding an entity as responsive simpliciter, we also often conclude that it is responsive to particular stimuli in its environment. In such cases, the entity is aware of or has transitive consciousness (Rosenthal 1986) of something. (This variety of consciousness is transitive because, like a transitive verb, it requires an object.) Responsiveness to one's environment presupposes transitive consciousness of aspects of one's environment. But transitive consciousness without responsiveness seems at least conceptually possible. We sometimes wonder if a comatose person is aware that his loved ones are nearby. In such cases, we wonder whether he has transitive consciousness, though he clearly lacks responsiveness. What's more, many consciousness researchers posit transitive awareness of one's own mental states in addition to the awareness of environment that responsiveness requires. Indeed, some go so far as to identify transitive consciousness of certain of one's own mental states with phenomenal consciousness (see, for example, Armstrong 1968; Rosenthal 1986; Lycan 1996).

Phenomenal Consciousness:

Many philosophers of mind contend that a number of mental states are alike in exhibiting their own distinctive experiential qualities. Mental states such as believing and desiring, on the other hand, are purported not to exhibit these qualities—so, not to participate in phenomenal consciousness—according to these philosophers. Though no one has yet satisfactorily defined phenomenal consciousness, Ned Block characterizes it by writing that, “what makes a state phenomenally conscious is that there is something ‘it is like’ (Nagel 1974) to be in that state” (230). Where definition fails, many thinkers simply point: the distinctive pang of hunger, the feeling of happiness at seeing a restaurant, the rich taste experience of eating pastrami. Many claim that phenomenal experience is as inescapable as it is ineffable. (Indeed, the introduction of the present article trades on the purported inescapability of phenomenal consciousness.) Others argue that ordinary people do not recognize the philosopher's class of phenomenally conscious mental states at all—despite the fact that philosophers such as David Chalmers (1995, p. 207) deem phenomenal consciousness “the

most central and manifest aspect of our mental lives.” Still others deny the existence of any rich variety of phenomenal consciousness outright (Dennett 1990). Regardless of whether some of our mental states actually have manifest, qualitative feels akin to the sorts of properties philosophers of mind have subsumed under phenomenal consciousness, a number of experimentalists have compiled evidence suggesting that ordinary people draw roughly the same dichotomy between phenomenal and intentional states that philosophers of mind have long alluded to (Gray et al. 2007; Knobe and Prinz 2008). Indeed, many of the papers in this issue build on or challenge this work, suggesting interconnections between moral psychology and phenomenal consciousness attribution, or challenging such connections posited by previous theorists.

Self-Consciousness:

Finally, we commonly refer to an entity as self-conscious when it can occupy any of a range of representational states that have as their content (states of) the entity itself. Most generally, an entity can be said to be self-conscious if it has a concept of itself as a particular entity, as is taken to be evidenced by Gordon Gallup Jr’s (1970) mirror test. In this test, sedated animals are marked with a smudge of paint. They are judged to have passed the test if, after they wake up, they investigate the spot on their body after examining their image in a mirror. (To date, all great apes, some cetaceans, elephants, and European Magpies have been shown to pass the test.) Richer states of self-consciousness include representations of one’s mental states as state’s of one’s mind, beliefs about oneself as an entity enduring over a span of time, and other ideas about one’s nature. These richer states involve various conceptual apparatus over and above the minimal cognitive capacity required to pass the mirror test (presumably, the capacity to identify the reflected body with an occurrent, though not necessarily enduring, self). Identification with one’s own mental states requires metacognition. Long-term self-conceptions require a concept of an enduring self. Though some attributions of self-consciousness seem to require self-knowledge (for example, varieties of self-awareness), we will not presuppose self-knowledge even for richer states of self-conception. Misconceptions of self must nonetheless represent the self; and so a narcissist who thinks himself humble is self-conscious, though he fails to know himself. Several papers in this issue focus on self-consciousness. These papers ask how do conceptions of one’s self or conceptions of others’ conceptions of themselves factor into moral behavior and assessment?

In addition to the four concepts of consciousness discussed above, there are a number of other things one might mean by the term consciousness. For example, Block (1995) identifies a special kind of transitive consciousness, *access conscious*, encompassing representations that are occurrently available to guide reasoning and action at the person-level. And, of course, we sometimes speak in a general sense, of a *kind* of organism as being *sentient*, or responsive to its world, in a way that does not require occurrent responsiveness of any particular organism as outlined above. A thorough study would reveal further uses. Nonetheless, those concepts of consciousness discussed above serve to outline some of the key aspects of consciousness with

which philosophers and psychologists have been concerned. They also include those notions of consciousness upon which authors in this volume focus.

Having surveyed the many things one may mean by the word “consciousness,” a question arises: Why do all these disparate states belong together, much less constitute a useful topic for a special issue of a journal? We believe that treating consciousness as a unified phenomenon for this special issue of *Review of Philosophy and Psychology* is warranted because these various “consciousnesses” are apparently tightly interrelated. Many are at least partially definable in terms of the others. As aforementioned, responsiveness to one’s environment seems to involve a variety of transitive awareness; whereas self-consciousness is transitive awareness of oneself and the properties thereof. Furthermore, where interdefinability is not obvious, various theorists have attempted to explain one or another of these concepts of consciousness in terms of others. Externalist representationalists, including Dretske (1995) and Tye (1995), attempt to explain phenomenal consciousness (the “what it’s likeness” of some mental states) in terms of a special variety of transitive consciousness of external stimuli. Higher-order theorists, such as Rosenthal (1993) and Lycan (1996) explain it in terms of a special variety of self-consciousness. Perhaps more importantly for our specific topic, treating consciousness as a unified phenomenon for purposes of investigating the relation between consciousness and moral cognition is warranted because it conforms to the practice of many of the most successful contemporary researchers into moral psychology. We review this research in the next section.

2 Moral Cognition

The study of moral cognition has its roots in developmental psychology, in which two landmark figures, Jean Piaget (1932) and Lawrence Kohlberg (1958) proposed stage models for how children come to reason about right and wrong. Although both models are best characterized as generalizations, both link sophisticated moral judgments of an actor to identifications of the actors’ mental states, specifically the actor’s intentions. This recognition that the perception of others’ conscious states influences moral reasoning paralleled early work in social psychology that began to explore the factors that led people to treat others in a moral or immoral fashion. Much of this work looked at the ways that the perception of others as fully conscious humans influenced aggressive or prosocial behavior. For example, much research suggested that dehumanization—treating others as lacking fundamental conscious states—licenses immoral and aggressive action toward others, and also served to justify these actions retroactively (Bandura et al. 1975; Kelman 1976). Similarly, research on deindividuation suggested that the lack of awareness of one’s *own* mental states can license immoral behavior as well, through diminishing feelings of personal responsibility (Milgram 1963; Zimbardo 1969).

As social, developmental, and cognitive psychologists continued to study moral cognition through the latter half of the 20th century, much progress was made on identifying additional factors that contributed to decisions to act morally or immorally and to judge others’ actions as right and wrong. Some of these factors again involved others’ mental states (Karniol 1978) whereas others ranged from the actions’

outcomes (Walster 1966), to culture (Shweder et al. 1987), to norms for justice (Turiel 1983). At the turn of the 20th century, Jonathan Haidt (2001) published a landmark paper suggesting that people do strikingly little conscious consideration of various factors when determining right and wrong; rather people make these judgments based on emotional, gut reactions. This seminal paper set off a flurry of research on morality and inspired a large body of research to investigate the extent to which moral cognition truly involves cold, conscious deliberation or is more emotional and intuitive in nature.

In this resurgence of morality research, again much discussion centered on the role of recognizing others' mental states in moral judgment and action. Some recent work has suggested that the recognition of others' mental states is of central importance to morality (Gray et al. 2007; Gray et al. 2012). This work states that judgments of whether an action is morally wrong center on whether the action involved an actor with thought and intention (a moral agent) acting toward an entity that experiences emotion and feeling as a result of the action (a moral patient). In other words, moral judgments of action center on whether the doer of the action has the mental capacity of agency and on whether the recipient of the action has the mental capacity of experience. Much work supports this distinction between the recognition of agency in moral agents and experience in moral patients (Gray and Wegner 2009; Robbins and Jack 2006), whereas other work challenges the extent to which agency and experience are processed separately in moral judgments (Arico et al. 2011). What is more, numerous prominent accounts of morality largely disregard the role of others' conscious states in moral judgments altogether. For example, moral foundations theory (Graham et al. 2009; Haidt 2007) suggests that at least six innate psychological principles guide moral judgments, and most of these principles—whether an action involved loyalty to one's group or whether an action was carried out in accord with sanctity—depend on recognizing others' mental states. Other emerging theories such as relationship regulation theory (Rai and Fiske 2011) similarly do not prioritize understanding of others' conscious states, but rather suggest that people base moral judgments of an action based on whether the action upholds or degrades specific types of social relationships. The present issue investigates the relationship between consciousness and moral cognition, exploring the extent to which perception of agency and experience influences moral judgment, the extent to which they operate separately or jointly, and the extent to which the recognition of these conscious states in others is at all necessary for moral cognition.

3 The Current Issue

Papers in this issue can be grouped along several dimensions. First, each of the present papers is fundamentally either a defense of an important psychological connection between consciousness (broadly construed) and moral cognition, or else a challenge to one or another cognitive theory of how consciousness and moral cognition are related. Amongst those papers that posit an important connection between consciousness and moral cognition, we can further delineate between three groups. Some of these papers emphasize consciousness and argue that either phenomenal state attribution or awareness of self or other is linked to moral cognition in

novel and distinctive ways. Other of these papers emphasize moral cognition and argue that certain patterns of moral cognition typically lead to consciousness attribution. Finally, some of the papers in this issue argue in favor of a tight, bi-directional coupling between consciousness and moral-cognition. Yet a third dimension cross-cuts those previously discussed, for while the majority of the papers in this issue focus on phenomenal consciousness, others primarily discuss self-awareness or awareness of others.

Papers arguing that consciousness is linked to moral cognition include Justin Sytsma and Edouard Machery's "The Two Sources of Moral Standing." Sytsma and Machery argue that experience and agency are independent cues for lay judgments regarding moral standing. Such lay tendencies, they maintain, have influenced philosophical theorizing, as is evidenced by the fact that classic ethical theories emphasize either rationality or experience. In "Doing good leads to more good: The reinforcing power of a moral self-concept," Liane Young, Alek Chakroff, and Jessica Tom, contend that having the right kind of self-conception can reinforce moral behavior. Specifically, Young et al. found that people primed to think of their own good deeds, yet disinclined to think of how others reacted to those deeds, gave to charity at higher rates than others. Finally, Adam Feltz and Edward T. Cokely's "The Virtues of Ignorance" argues that erroneous self-conceptions are sometimes more virtuous than accurate ones.

The previous papers are alike in suggesting that if one exhibits or recognizes consciousness (in one sense of the term or another), then one will be disposed to behave or cognize in some morally relevant way. A second set of papers in this issue come at the topic from the opposite direction. These papers suggest that if one deploys certain moral concepts or has certain proto-moral needs, then one will be inclined to ascribe conscious states to others. In "Explaining the Abstract/Concrete Paradoxes in Moral Psychology: The NBAR Hypothesis," Eric Mandelbaum and David Ripley make the case for a view according to which ordinary people have an unconscious belief that when a norm (moral or otherwise) is broken, an agent is responsible. When no obvious agentic candidate is available, this unconscious belief leads people to attribute agency to a non-obvious candidate (e.g., an immaterial spirit, God, or an inanimate object). Ida Halgren, on the other hand, argues for a bifurcated view of mental state attribution in "Seeing Agents When We Need To, Attributing Experience When We Feel Like It." She distinguishes between a "doing-mode" of action—in which it may be important to attribute intentional states, though not experiences—and an "interactive mode"—in which understanding the experiential states of others is key.

Turning from papers that emphasize one direction of the bi-conditional or the other, we come to those that argue for a bi-directional coupling between consciousness and moral-cognition. Anthony I. Jack and Philip Robbins revisit their phenomenal stance hypothesis, from their (2006) paper. Here they refine their view, according to which we can adopt a distinct phenomenal stance for interpreting the experiential states of others. They offer empirical evidence for this view, as well as for a bi-directional connection between consciousness and moral cognition and for a subtle tension in regarding something as both a physical and a phenomenal entity. In their paper, "Two Minds Versus Two Philosophies: Mind Perception Defines Morality and Dissolves the Debate Between Deontology and Utilitarianism," Kurt Gray and

Chelsea Schein argue that strict forms of both deontology and consequentialism are psychologically untenable because moral cognition is essentially dyadic. According to their moral typecasting view, “moral judgments are based on a fundamental template of two perceived minds—an intentional agent and a suffering patient.” Perceiving an agent invariably draws the mind to search for a patient, and vice versa. Thus we cannot focus on either an agent or a patient in isolation, as, Gray and Schein argue, strict forms of deontology and consequentialism require.

The final set of papers in this volume challenge one or another aspect of the moral cognition/consciousness connection. In “Breaking Out of Moral Typecasting,” Adam Arico takes on the body of empirical research on which the moral typecasting view is based, arguing that a small set of confounds do a better job of explaining the data in a more parsimonious fashion. Deena Skolnick Weisberg and Alan M. Leslie’s, “The Role of Victim’s Emotions In Preschoolers Moral Judgments,” serves to partially uncouple consciousness perception and moral cognition by presenting evidence that “typically developing children’s moral judgments are...concerned primarily with action type, not with emotional displays.” And Brent Strickland and Aysu Suben make the case that some key evidence in favor of one popular account of the relation between consciousness and moral cognition may be based on a subtle form of experimenter bias, in their “Experimenter philosophy: The problem of experimenter bias in experimental philosophy.”

Much remains to be written about the relation between consciousness and moral cognition. We hope that the present contributions will inspire other researchers to take up this exciting topic.

References

- Arico, A., B. Fiala, R.F. Goldberg, and S. Nichols. 2011. The folk psychology of consciousness. *Mind & Language* 26(3): 327–352.
- Armstrong, D. 1968. *A materialist theory of mind*. London: Routledge and Kegan Paul.
- Bandura, A., B. Underwood, and M.E. Fromson. 1975. Disinhibition of aggression through diffusion of responsibility and dehumanization of victims. *Journal of Research in Personality* 9: 253–269.
- Blair, R.J.R., C. Sellars, I. Strickland, F. Clark, A.O. Williams, M. Smith, and L. Jones. 1995. Emotion attributions in the psychopath. *Personality and Individual Differences* 19: 431–437.
- Block, N. 1995. On a confusion about the function of consciousness. *Behavioral and Brain Sciences* 18: 227–247.
- Chalmers, D. 1995. Facing up to the problem of consciousness. *Journal of Consciousness Studies* 2: 200–219.
- Dennett, D.C. 1990. Quining qualia. In *Mind and cognition*, ed. W. Lycan, 519–548. Oxford: Blackwell.
- Dretske, F. 1995. *Naturalizing the mind*. Cambridge: The MIT Press, Bradford Books.
- Gallup Jr., G.G. 1970. Chimpanzees: self-recognition. *Science* 167: 86–87.
- Graham, J., J. Haidt, and B.A. Nosek. 2009. Liberals and conservatives rely on different sets of moral foundations. *Journal of Personality and Social Psychology* 96: 1029–1046.
- Gray, H.M., K. Gray, and D.M. Wegner. 2007. *Dimensions of mind perception*. *Science* 315: 619.
- Gray, K., and D.M. Wegner. 2009. Moral typecasting: divergent perceptions of moral agents and moral patients. *Journal of Personality and Social Psychology* 96: 505–520.
- Gray, K., L. Young, and A. Waytz. 2012. Mind perception is the essence of morality. *Psychological Inquiry* 23: 101–124.
- Haidt, J. 2001. The emotional dog and its rational tail: a social intuitionist approach to moral judgment. *Psychological Review* 108: 814–834.
- Haidt, J. 2007. The new synthesis in moral psychology. *Science* 316: 998–1002.

- Kant, I. 1785/1998. *Groundwork of the metaphysics of morals*, trans. M. Gregor. Cambridge: Cambridge University Press.
- Karniol, R. 1978. Children's use of intention cues in evaluating behavior. *Psychological Bulletin* 85: 76–85.
- Kelman, H.C. 1976. Violence without restraint: Reflections on the dehumanization of victims and victimizers. In *Varieties of psychohistory*, ed. G.M. Kren and L.H. Rappoport, 282–314. New York: Springer.
- Knobe, J., and J. Prinz. 2008. Intuitions about consciousness: experimental studies. *Phenomenology and Cognitive Science* 7: 67–83.
- Kohlberg, L. 1958. *The development of modes of moral thinking and choice in the years 10 to 16*. Doctoral dissertation, University of Chicago.
- Lycan, W. 1996. *Consciousness and experience*. Cambridge: MIT Press.
- Milgram, S. 1963. Behavioral study of obedience. *Journal of Abnormal and Social Psychology*. 67: 371–378.
- Miller, D.T., and M. Ross. 1975. Self-serving biases in the attribution of causality: fact or fiction? *Psychological Bulletin* 82: 213–225.
- Nagel, T. 1974. What is it like to be a bat? *Philosophical Review* 83: 435–450.
- Piaget, J. 1932. *The moral judgment of the child*. New York: Harcourt Brace.
- Rai, T.S., and A.P. Fiske. 2011. Moral psychology is relationship regulation: moral motives for unity, hierarchy, equality, and proportionality. *Psychological Review* 118: 57–75.
- Robbins, P., and A.I. Jack. 2006. The phenomenal stance. *Philosophical Studies* 127(1): 59–85.
- Rosenthal, D.M. 1986. Two concepts of consciousness. *Philosophical Studies* 49(May): 329–359.
- Rosenthal, D.M. 1993. Thinking that one thinks. In *Consciousness: Psychological and philosophical essays*, ed. M. Davies and G. Humphreys. Oxford: Blackwell.
- Shweder, R.A., M. Mahapatra, and J. Miller. 1987. Culture and moral development. In *The emergence of morality in young children*, ed. J. Kagan and S. Lamb, 1–83. Chicago: University of Chicago Press.
- Singer, P. 1975. *Animal liberation: A new ethics for our treatment of animals*. New York: New York Review/Random House.
- Turiel, E. 1983. *The development of social knowledge: Morality and convention*. Cambridge: Cambridge University Press.
- Tye, M. 1995. *Ten problems of consciousness*. Cambridge: MIT Press.
- Van Gulick, R. 2011. "Consciousness", *The Stanford Encyclopedia of Philosophy (Summer 2011 Edition)*, Edward N. Zalta (ed.), <<http://plato.stanford.edu/archives/sum2011/entries/consciousness/>>.
- Walster, E. 1966. Assignment of responsibility for an accident. *Journal of Personality and Social Psychology* 3: 73–79.
- Wenar, L. 2010. "Rights." *The Stanford Encyclopedia of Philosophy (Fall 2010 Edition)*, Edward N. Zalta (ed.), <<http://plato.stanford.edu/archives/fall2010/entries/rights/>>.
- Zimbardo, P.G. 1969. The human choice: Individuation, reason, and order versus deindividuation, impulse, and chaos. In *1969 Nebraska symposium on motivation*, ed. W.J. Arnold and D. Levine, 237–307. Lincoln: University of Nebraska Press.