"OF WHAT USE ARE THE *ODES*?" COGNITIVE SCIENCE, VIRTUE ETHICS, AND EARLY CONFUCIAN ETHICS

Edward Slingerland
Canada Research Chair in Chinese Thought and Embodied Cognition; Department of Asian Studies, University of British Columbia

In his well-known 1994 work *Descartes' Error,* the neuroscientist Antonio Damasio describes his work with patients suffering from damage to the prefrontal cortex, a center of emotion processing in the brain. The accidents or strokes that had caused this damage had spared these patients' "higher" cognitive faculties: their short- and long-term memories, abstract reasoning skills, mathematical aptitude, and performance on standard IQ tests were completely unimpaired. They were also perfectly healthy physically, with no apparent motor or sensory disabilities. Nonetheless, these patients had been brought to Damasio's attention as a physician because, despite their apparent lack of physical or cognitive impairment, they were no longer functional members of society. In real-life decision-making contexts they were appallingly inept, apparently incapable of efficiently choosing between alternate courses of action, taking into account the future consequences of their actions, or accurately prioritizing the relative importance of potential courses of action.

One representative example is the patient Damasio refers to as "Elliot." Formerly a successful businessman and respected husband and father, Elliot's life began to unravel after he was operated on for a brain tumor, a procedure that involved removing parts of his prefrontal cortex. Elliot needed to be prompted to get up for work in the morning, and once there seemed incapable of managing his time properly, focusing his attention effectively, or completing even the most routine of tasks. Understandably, Elliot was soon fired. He proved no more successful in negotiating his way through unemployed life. He developed bizarre collecting habits, took up a bewilderingly diverse array of temporary projects, entered into questionable financial ventures with disreputable individuals, lost his life's savings, divorced, briefly remarried to a woman of questionable character, divorced again, and finally—completely destitute and without any means of support—was reduced to living off of social security disability payments. Elliot's behavioral profile is fairly typical of individuals with prefrontal cortex damage. Damasio describes a similar inability to contextualize actions within a larger framework of meaning in an anecdote concerning an attempt to set up a follow-up appointment with another prefrontal patient of his. This individual pulled out his appointment book and launched into a tiresome cost-benefit analysis, spending approximately thirty minutes enumerating the reasons for and against two alternate dates, citing any factor that could conceivably impinge on the decision—previous engagements, possible meteorological conditions, et cetera—until an exasperated Damasio finally intervened and simply decided for him (Damasio 1994, pp. 194–195).

Philosophy East & West Volume 61, Number 1 January 2011 80–109
© 2011 by University of Hawai'i Press

(ID CS4 6 September 2010 12:49 PM) UHP (7×10") Optima J-2344 PEW 61:1 (idp) PMU:(idp[KN]/A)6/9/2010 pp. 80–109 2344_61-1_04 (p. 80)

One of the things that I found most interesting about Elliot's case is that Elliot, in addition to performing very well on general IQ tests and other measures of abstract cognitive ability, also scored quite well on the Standard Issue Moral Judgment Interview, developed by the Kantian moral psychologist Lawrence Kohlberg, which measures a person's ability to reason their way abstractly through moral dilemmas and other theoretical problems. Apparently, this theoretical ability to reason about dilemmas did not translate into an ability to make actual reasonable ethical *decisions*: "at the end of one session, after he had produced an abundant quantity of options for action, all of which were valid and implementable, Elliot smiled, apparently satisfied with his rich imagination, but added, 'And after all this, I still wouldn't know what to do!'" (Damasio 1994, p. 49).

When I first read about Elliot's predicament—that is, having elaborate theoretical knowledge about what he should do, but being completely unable to actually decide what *to* do—I was immediately reminded of *Analects* 13.5, where Confucius remarks, "Imagine a person who can recite the several hundred Odes by heart but, when delegated a governmental task, is unable to carry it out, or when sent abroad as an envoy, is unable to engage in repartee. No matter how many Odes he might have memorized, of what use are they to him?" As many commentators have noted, the point of Confucius' remark is that learning involves not merely the acquisition of abstract scholastic knowledge, but more importantly the ability to flexibly apply this knowledge in real-life situations. As is made quite clear throughout early Confucian texts, this sort of ability is only acquired once one has learned how to *feel* the right sort of things at the right sorts of times, and this sort of feeling can only be produced by an acculturation process—involving music, ritual, and role-model emulation—that transforms both the body and the mind of the student.

This so-called "virtue ethics" model of moral self-cultivation differs in important respects from the deontological and utilitarian models that currently dominate modern Western ethical discourse that are based on a disembodied, rationalist model of the self.[1] For instance, in a rule-based deontological approach, we have certain maxims, such as "It is wrong to lie." When presented with a situation, we can consult our definition of a "lie" to determine whether act X in this given situation was or was not an instance of lying, and then once this is determined we can decide whether it was right or wrong depending on where this particular maxim is located in a hierarchy of maxims—for example, perhaps it is trumped by the maxim that we should strive to preserve life. If we are utilitarians, in any situation we should be able unproblematically to tally up the costs and benefits of proposed courses of action, do the math, and thereby figure out which course of action maximizes whatever good our brand of utilitarianism deems important (happiness, justice, gross national product, etc.). In either case, the entire process of moral reasoning is transparent and under our conscious control, and has nothing to do with the details of our embodiment, or with emotions, implicit skills, or unconscious habits.[2]

What I intend to do below is review recent evidence from the fields of cognitive science, cognitive linguistics, behavioral neuroscience, and social psychology that call into question this model of the self and that point to the crucial role of affect,

(ID CS4 6 September 2010 12:49 PM) UHP (7×10") Optima J-2344 PEW 61:1 (idp) PMU:(idp[KN]/A)6/9/2010 pp. 80–109 2344_61-1_04 (p. 81)

embodiment, and metaphorical extension in moral judgments and decision making. I will then examine the implications of these results for moral philosophy, concluding that the virtue-ethical model of self-cultivation more accurately represents how real human beings engage in moral reasoning and is also better adapted as an educational technique to the evolved cognitive architecture of human beings than deontological or utilitarian approaches. Finally, turning to early Confucianism, I will suggest some of the ways in which Mencius' views about morality and ethical education strikingly anticipate findings coming out of the modern Western cognitive sciences, and therefore how thinkers such as Mencius can serve as an important conceptual resource in envisioning what an empirically responsible modern virtue ethic might look like.

Some Themes Emerging from Cognitive Science[3]

Thought Is Image-Based

One of the most fundamental challenges to the disembodied, amodal model of human reason is the increasing consensus in the fields of neuroscience and cognitive science that human thought is primarily image-based and modal in character—that is, deriving its structure from sensory-motor patterns. As opposed to a picture of thought as the manipulation of arbitrary, abstract symbols, cognitive scientists such as Lawrence Barsalou have been arguing for a "perceptual symbol" account of human cognition.[4] According to this model, the symbols manipulated in human thought are understood not as pictures but as "records of neural activation that arises during perception" (Barsalou 1999, p. 583). These records can be abstracted from and combined in various ways in areas of the brain "upstream" from the sensory-motor cortices (what Damasio [1989] refers to as "convergence zones"), but they always remain to some extent grounded in sensory-motor systems.

There is a huge and constantly growing body of evidence in favor of at least some version of the perceptual-symbol account, including evidence that conceptual symbols have inextricable perceptual character, that the affordances derived from sensory-motor simulations are essential to semantic processing, and that damage to sensory-motor systems results in category-specific deficits in cognition. For instance, damage to visual areas selectively disrupts the conceptual processing of categories specified by visual features (e.g., birds), while damage to the motor regions selectively disrupts use of categories specified by motor programs (e.g., tools).[5] Work on imitation has found that both the perception and conceptualization of action and action-related words requires the activation of the appropriate sensory-motor regions of the brain (e.g., Rizzolatti et al. 2001). Perhaps the strongest argument in favor of something like the perceptual-symbol account is that it avoids two fundamental problems that plague amodal symbolic accounts: the transduction problem (how perceptual signals could get "translated" into amodal symbols) and the grounding problem (how arbitrary, abstract symbols could ever come to refer to something in the world). As Barsalou concludes in his argument against classical amodal theories of meaning, such theories "are unfalsifiable, they are not parsimonious, they lack

(ID CS4 6 September 2010 12:49 PM) UHP (7×10") Optima J-2344 PEW 61:1 (idp) PMU:(idp[KN]/A)6/9/2010 pp. 80–109 2344_61-1_04 (p. 82)

direct support, they suffer conceptual problems such as transduction and symbol grounding, and it is not clear how to integrate them with theory in neighboring fields, such as perception and neuroscience" (Barsalou 1999, p. 580).

Categories are Usually Radial and Based on Prototypes

The rationalist approach to reasoning relies on classic Aristotelian categories, which have sharp boundaries and clearly defined sufficient and necessary conditions for category membership. Evidence from cognitive psychology and linguistics suggests that the mode of categorization generally relied on by human beings differs significantly from the classical account. Much of the early work in this field was done by Eleanor Rosch and her colleagues,[6] who developed a theory of "radial" categorization based on a "prototype effect." Categories as they are usually active in human minds are based on certain exemplars or prototypes; membership in the category is then based on family resemblance and can be a matter of degree (there can be "better" or "worse" members of a given category). For example, most North Americans have an understanding of the category "bird" that is based on an image of a sparrow, robin, or jay. Most people can switch into a "logical category" mode and acknowledge that chickens, penguins, and ostriches are "birds," but will continue to insist that these are not particularly "good" examples of birds. The same effect can be seen with social categories such as "bachelor": the Pope, for instance, is not a particularly good instance of a "bachelor" (Lakoff 1987).

The dominance in everyday thought of prototype-based categorization is expected from the perspective of the perceptual-symbol account of cognition. If concepts are a form of sensory-motor simulation, categorization will be based on imagined exemplars and family resemblances. As Barsalou notes, categorization understood from this perspective will also not be designed as a rigid net for exhaustively cataloging and organizing sets of clearly defined objects in the world, but rather as a dynamic, contextual, and embodied means of gaining access to categorical inferences—that is, suggestions as to how to interact successfully with encountered objects and situations and to reason about absent (future) entities (Barsalou 1999, p. 587).

Judgment / Decision Making Is Grounded in Somatic-Sensory Emotional Reactions

In the last decade there has been an explosion of literature on the role of emotions in human reasoning in such fields as behavioral neuroscience, cognitive science, economics, social psychology, and philosophy.[7] Because of space restrictions, I will continue to focus on Damasio, and particularly his theory of "somatic marking." In his discussion of the "body-minded brain," Damasio points out that the mind evolved in order to assure the survival of the entire mind-body unit, and argues that the best way to do this is by "*representing the outside world in terms of the modifications it causes in the body proper,* that is, representing the environment by modifying the primordial representations of the body proper whenever an interaction between organism and environment takes place" (Damasio 1994, p. 230). The result is a set of "somatomotor maps" that provide a "dynamic map of the overall organism anchored in body

(Id CS4 6 September 2010 12:49 PM) UHP (7×10") Optima J-2344 PEW 61:1 (idp) PMU:(idp[KN]/A)6/9/2010 pp. 80–109 2344_61-1_04 (p. 83)

schema and body boundary" (p. 231). So, when we are presented with a situation—or called upon to imagine a situation (neurophysiologically not that different a process)—we rely on the "dispositional representations" (p. 104) that constitute our full repository of knowledge in order to comprehend it, and these representations inevitably include emotional information. As Damasio observes, "When we recall an object . . . we retrieve not just sensory data but also accompanying motor and emotional data. . . . We recall not just sensory characteristics of an actual object but the past reactions of the organism to the object" (p. 161). In other words, the images that form the basis of our concepts are somatically "marked" with visceral and often unconscious feelings of "goodness" or "badness," urgency or lack of urgency, and these feelings play a crucial role in everyday, "rational" decision making.

In the view of Damasio and his colleagues, the problem with prefrontal cortex patients such as "Elliot" is that they lack these somatic markers, which prevents them from unconsciously assigning different values to different options, thereby rendering their "decision-making landscape hopelessly flat" (Damasio 1994, p. 51). In any given situation, the number of theoretically possible courses of action is effectively infinite, and the human mind is obviously not capable of running simultaneous analyses of all of them at once. Therefore, the body contributes by biasing the reasoning process—often unconsciously—before it even begins. Patients such as Elliot perform well on abstract moral reasoning and utilitarian calculation tasks because such abstract analyses are artificially simplified. Thrown into a real-life situation, but deprived of the biasing function of somatic markers, they seem to attempt to dispassionately consider *all* of the options theoretically open to them, with the result that they become paralyzed by indecision or simply commit themselves to what appear to outside observers as poorly considered and capriciously selected courses of action.

Now, somatic markers do not necessarily always lead to advantageous decision making. The same patient of Damasio who was infuriatingly incapable of deciding on a date for the next appointment had calmly steered his way through a skid on icy roads earlier that same day—one sort of scenario where a person's immediate emotional response to a perceived danger typically leads to unhelpful behavior (slamming on the brakes). More generally, it is clear that human beings are often *bad*—that is, not rationally ideal—decision makers, especially when operating in modern industrial societies, far outside their ancestral environment.[8] Dispassionate calculation makes it clear that we are likely to achieve a much better payoff by investing twenty dollars weekly in some conservative mutual fund rather than using that money to buy lottery tickets, but the reasoning processes of many are (incorrectly, in this case) biased by the powerfully positive somatic marker attached to the image of the multimillion-dollar payoff. Similarly, the vividly negative image of a jetliner falling in flames from the sky prevents many from making the "rational" decision to fly rather than drive, even though commercial airline travel is demonstrably much safer than automobile travel.

George Loewenstein and his colleagues have formulated a "risks as feelings" hypothesis very similar to Damasio's somatic marker theory, finding that human risk

(ID CS4 6 September 2010 12:49 PM) UHP (7×10") Optima J-2344 PEW 61:1 (idp) PMU:(idp[KN]/A)6/9/2010 pp. 80–109 2344_61-1_04 (p. 84)

assessment of an imagined scenario is driven largely by vividness, not by the probability of that scenario actually occurring. In one study (Loewenstein et al. 2001) it was found that people are willing to pay more for airline travel insurance covering death from "terrorist acts" than for insurance covering death from "all possible causes"! At the other extreme, people tend to be underinsured against emotionally "pallid" risks such as floods. Other studies have found that people are also much more responsive to warnings that are linked to people and anecdotes than those put in statistical terms. Thus, while navigating by means of powerful, reasoning-biasing somatic markers must have been adaptive in our dispersed, hunter-gatherer "environment of evolutionary adaptation" (EEA), it sometimes leads us into errors of judgment in the more complex world of settled agricultural societies, especially when modern technology is thrown into the mix.

Judgment / Decision Making is Often Automatic and Unconscious

In their review of the social psychology literature on "automaticity," John Bargh and Tanya Chartrand discuss studies revealing the power of priming to affect modes of behavior, stereotype priming to affect social judgments, unconscious mimicry of behavior and its effect on social judgments (especially the case of mimicry inducing increased liking and sense of smooth interaction), and the unconscious acquisition of goals from external stimulation. They conclude that, with regard to many areas of their experience, people "classify their experience as either good or bad and do so immediately, unintentionally, and without awareness that they are doing it" (Bargh and Chartrand 1999, p. 474).

> Automatic evaluation of the environment is a pervasive and continuous activity that individuals do not intend to engage in and of which they are largely unaware. It appears to have real and functional consequences, creating behavioral readiness within fractions of a second to approach positive and avoid negative objects, and, through its effect on mood, serving as a signaling system for the overall safety versus danger of one's current environment. All of these effects tend to keep us in touch with the realities of our world in a way that bypasses the limitations of conscious self-regulation capabilities. (pp. 475–476)

In other words, the social psychology literature documents the pervasive importance of unconscious somatic markers on human behavior and attitude formation.

It is clear that there are separate systems that work on the implicit and explicit levels, with somatic markers functioning primarily at the former level. Robert Zajonc's famous studies demonstrated that people can have affective responses to stimuli before consciously knowing what it is that they are reacting to (Zajonc 1980, Kunst-Wilson and Zajonc 1980), and Damasio has shown that skin conductance reactions to emotionally charged stimuli *precede* conscious awareness of emotion: emotional states happen first, and conscious feelings follow (Damasio 2003, p. 101). Joseph Le Doux has postulated the existence of two systems of memory, an unconscious, implicit "emotional memory" and an explicit "declarative" memory (Le Doux 1996, pp. 181–182), and has reviewed studies indicating that priming, manual skills, and cog-

(Id CS4 6 September 2010 12:49 PM) UHP (7×10") Optima J-2344 PEW 61:1 (idp) PMU:(idp[KN]/A)6/9/2010 pp. 80–109 2344_61-1_04 (p. 85)

nitive skills (such as the ability to solve a particular type of puzzle) are preserved in amnesiac patients, suggesting that implicit "know-how" is developed and stored in brain systems separate from those that subserve conscious memory (pp. 195–198).

Of course, it is obvious that the brain systems associated with abstract reasoning and cognitive control can, at least sometimes, bring these implicit biases and other sorts of emotions into consciousness in order to modify or override them. Indeed, there is evidence that cortical control is necessary for the normal conscious experience and expression of emotion. Decorticate animals, for instance, have emotional reactions, but they are not entirely normal—such creatures are easily provoked and seem entirely incapable of regulating their emotional reactions, which suggests that cortical areas normally rein in and control emotional reactions (Le Doux 1996, p. 80). But it is equally clear that conscious self-control is something of a limited resource. The work of Roy Baumeister and his colleagues (Baumeister et al. 1998, Muraven et al. 1998) has shown that when conscious control is exerted in one domain, this depletes the individual's ability to exert it in another, unrelated domain. This suggests that conscious self-control must be a relatively rare occurrence, since it seems to require a lion's share of cognitive resources. There is also considerable evidence that conscious intervention in automatic processes can be counterproductive. Baumeister's work has shown that automatic behaviors are disrupted when people analyze and decompose them (Baumeister 1984). Similarly, Timothy Wilson and Jonathan Schooler have shown in a series of studies that, in many domains, people form automatic and apparently quite adaptive evaluations that can then be disrupted when these people are asked to reflect on their reasons for their evaluative feelings (Timothy Wilson et al. 1989, Timothy Wilson and Schooler 1991).

In summary, evolution seems to have off-loaded the vast bulk of our everyday decision making and judgment formation onto automatic, unconscious systems, because such systems are fast, frugal, and reliable.

There is No Unitary Self in Charge

The disembodied model of rational decision making requires a unitary, conscious "self," the locus of rationality and will, to serve as the maxim follower or utilitarian calculator, and also to enforce the decisions arrived at through moral reasoning upon the recalcitrant emotions or body. While this rational self is not always successful in exerting control over other portions of the self, it is at least *aware* of what it is doing and why.

The phenomenon of automaticity discussed above calls this assumption into question, as does the outline of the human neural architecture emerging from neuroscientific research. One of the main Cartesian "errors" that Damasio takes aim at in his 1994 book is the concept of a "Cartesian theater": a central area of consciousness that experiences the world and the self in a unified fashion and serves as a kind of headquarters of the self. As Damasio notes, there is no single region in the human brain equipped to act as such a central theater; although there are various intermediate-level "convergence zones" that coordinate information coming in from more specialized sensory-motor regions, there is no "master" convergence zone that has an

(ID CS4 6 September 2010 12:49 PM) UHP (7×10") Optima J-2344 PEW 61:1 (idp) PMU:(idp[KN]/A)6/9/2010 pp. 80–109 2344_61-1_04 (p. 86)

overall view of the entire process (Damasio 1994, pp. 94–96; also see Damasio 1989). Of course, in our everyday experience we *feel* a strong sense of mental integration—the intuition of a unified self in charge of and informed about everything is very powerful and universal—but this is, Damasio argues, "a trick of timing," an illusion "created from the concerted action of large-scale systems by synchronizing sets of neural activity in separate brain regions" (1994, p. 95). How this sort of "binding" occurs is still not precisely understood, but what is clear is that there is no little homunculus running a central command post in the brain.

Among the more dramatic illustrations of the de-centered nature of the self are the studies of split-brain patients performed by Michael Gazzaniga and his colleagues. In these patients the corpus callosum, which normally connects the left and right hemispheres, has been severed (this has been found to be an effective, if last-resort, treatment for certain severe forms of epilepsy). The left brain is the seat of verbal ability and interpretative synthesis—in other words, the locus of our sense of unified self—and Gazzaniga and his colleagues found that the illusion of an in-control, unified self that the left hemisphere weaves persists even when it is most certainly *not* in control. For instance, in one experiment, images were selectively presented to each hemisphere: the left hemisphere was shown a chicken claw, the right a snow scene. Subjects were then presented with an array of objects and asked to choose an object "associated" with the image they were shown. A representative response was that of a patient who chose a snow shovel with his left hand (controlled, as we know, by the right hemisphere) and a chicken with the right (controlled by the left hemisphere). Asked why he chose these items, "he" (that is, his left hemisphere "spin doctor") replied, "Oh, that's simple. The chicken claw goes with the chicken, and you need a shovel to clean out the chicken shed" (Gazzaniga 1998, p. 25). Gazzaniga and Le Doux found a similar effect with normative judgments: in one particular patient referred to as "P.S.," the left hemisphere could correctly identify the emotional valence of a stimulus presented to the right hemisphere ("good" or "bad") without any conscious awareness of what the stimulus was (Le Doux 1996, pp. 14–15). In other words, the left hemisphere "was making emotional judgments without knowing what was being judged" (p. 15).

As Gazzaniga concludes, "The left brain weaves its story in order to convince itself and you that it is in full control" (Gazzaniga 1998, p. 25). He argues that, in place of the all-powerful legislator or canny calculator, a more appropriate metaphor for the conscious, verbal self might be a "harried playground monitor, a hapless entity charged with the responsibility of keeping track of multitudinous brain impulses running in all directions at once" (p. 23), and also responsible for concocting an ex post facto story of unified control for the consumption of both itself and others. One is reminded of Nietzsche's claim that the idea of free will is "the expression for the complex state of delight of the person exercising volition, who commands and at the same time identifies himself with the executor of the order," taking pleasure in the illusion that *"L'effet c'est moi"* (Nietzsche 1886/1966, p. 26).

Lest one think this sort of illusion of self-control is confined to people with extreme trauma, such as a severed corpus callosum, a large body of psychological ex-

(ID CS4 6 September 2010 12:49 PM) UHP (7×10") Optima J-2344 PEW 61:1 (idp) PMU:(idp[KN]/A)6/9/2010 pp. 80–109 2344_61-1_04 (p. 87)

perimental evidence has demonstrated the existence of a completely deluded "spin doctor" in neurologically "normal" individuals. The classic experiments in this field were performed by Richard Nisbett and Timothy Wilson (Nisbett and Wilson 1977, Wilson and Nisbett 1978), who demonstrated that people often report having thoughts and desires that they do not, in fact, have, and that the verbal reports given by subjects concerning the effects of stimuli on their judgments and behavior in experiments are often highly inaccurate. Studies of subjects given posthypnotic suggestions show a similar effect. For instance, Philip Zimbardo et al. (1993) found that subjects in which both hypnotic arousal and amnesia were induced generated a range of plausible explanations for their mental state that had nothing to do with the actual context of the experiment, and Paul Rozin and Carol Nemeroff (1990) found that subjects justified disgust-based attitudes with rationalizations that proved, upon examination, to be poor predictors of their actual behavior. Together with the vast literature on the unconscious effects of stereotype, mood, and emotional priming,[9] these results suggest that the deontological or utilitarian self is not, in fact, master of its own house.

Spelling Out the Implications for Moral Philosophy

Moral Reasoning Is Imaginative, Prototype-Based, and Non-Algorithmic

That imagination is crucial for moral reasoning has been the central argument of Mark Johnson for two decades (see especially Johnson 1987 and 1993), and has also been argued by philosophers such as Martha Nussbaum. In a discussion of the importance of imagination and literature for morality, Nussbaum notes that moral knowledge "is not simply intellectual grasp of propositions; it is not even simply intellectual grasp of particular facts; it is perception. It is seeing a complex, concrete reality in a highly lucid and richly responsive way; it is taking in what is there, with imagination and feeling" (Nussbaum 1990, p. 152). Part of what this sort of moral perception involves is the categorization of novel situations in terms of learned prototypes, which in turn involves a kind of intuitive pattern-matching rather than conscious rule-following. Johnson discusses the work of Linda Coleman and Paul Kay (1981) (refined by Eve Sweetser [1987]) on the prototype semantics of the English word "lie," which seems to demonstrate a radial category structure: there are better and worse instances of what constitutes a "lie," and subjects' judgment of whether or not a given act constitutes a lie depends on a set of implicit criteria that are contextually weighted, as well as on what Sweetser refers to as "idealized cognitive models" of knowledge and communication.[10] Applying these models to novel situations involves the re-activation of previous sensory-motor experiences, the identification of relevant features in the novel situation, and the recruitment of both implicit and explicit social knowledge. This process cannot be captured in an algorithmic maxim-following or cost-benefit analysis.

What this means is that moral education will involve training individuals—explicitly or implicitly—to develop more and more sophisticated imagistic models,

(ID CS4 6 September 2010 12:49 PM) UHP (7×10") Optima J-2344 PEW 61:1 (idp) PMU:(idp[KN]/A)6/9/2010 pp. 80–109 2344_61-1_04 (p. 88)

as well as the ability to extend them in a consistent manner. As Johnson explains, in any kind of reasonably complex situation "moral reasoning cannot consist merely in the rational unpacking of a determinate concept. Instead, it requires imaginative extensions to nonprototypical cases" (Johnson 1993, p. 100). Such extensions often involve the use of metaphors or analogies, and thus both internal moral reasoning and public moral debate will often take the form of battling metaphors—which metaphor or analogy best captures the current situation?[11] Is the current U.S. position in Iraq a "quagmire" like Vietnam, or is it like the difficulties encountered in the early period of implementing the Marshall Plan? When a senator vetoes an aid bill to help Sudanese famine victims, is he snatching food out of the mouths of hungry children, or is he helping the Sudanese to learn to stand on their own two feet? Are people seen leaving a ruined supermarket in flooded New Orleans "looters," or are they scrappy "survivors" making the best of a bad situation? How we choose to metaphorically frame a situation is probably the single most crucial element in how we will morally reason and morally *feel* about it, which leads us to our second point.

Moral Evaluations are Based on Emotions

A growing number of cognitive scientists and philosophers have come to agree with David Hume and the Stoics that normative judgments are ultimately derived from human emotional reactions. Damasio has argued that our sense of "goodness" corresponds to our sense of bodily wellness, which is not surprising considering that "achieving survival coincides with the ultimate reduction of unpleasant bodily states and the attaining of homeostatic ones, i.e., functionally balanced biological states" (Damasio 1994, p. 179). Martha Nussbaum has similarly argued for a "neo-Stoic" "cognitive-evaluative" view of emotions, which views them as "intelligent responses to the perception of value" (Nussbaum 2001, p. 1).[12]

The work of Charles Taylor (1989) is helpful in this regard. One of Taylor's most important points is that human beings, by their very nature, can only operate within the context of a normative space defined by a framework of empirically unverifiable beliefs. The Enlightenment conceit that one can dispense with belief or faith entirely, and make one's way through life guided solely by the dictates of objective reason, is nothing more than that: a conceit, itself a type of faith in the power of a mysterious faculty, "reason," to reveal incorrigible truth. In addition to the panoply of "weak evaluations"—such as a preference for chocolate over vanilla ice cream—that we are familiar with, humans are also inevitably moved to assert "strong" or normative evaluations. This latter type of evaluation is based on a set of explicit or implicit ontological claims, and therefore is perceived as having objective force rather than being a merely subjective whim. For instance, I don't particularly like chocolate ice cream, and believe that the flavor of vanilla ice cream is superior. I don't, however, expect everyone to share my preference, and am certainly not moved to condemn my wife for preferring chocolate. I am also not inclined to sexually abuse small children, but this feels like a different sort of preference to me: abusing small children seems *wrong,* and I would condemn and be moved to punish anyone who acted in

(Id CS4 6 September 2010 12:49 PM) UHP (7×10") Optima J-2344 PEW 61:1 (idp) PMU:(idp[KN]/A)6/9/2010 pp. 80–109 2344_61-1_04 (p. 89)

a manner that violated this feeling. If I were pressed on the matter, this condemnation would be framed, moreover, in terms of beliefs about the value of undamaged human personhood and the need to prevent suffering and safeguard innocence.

In the field of cognitive science and social psychology there is a growing set of empirical literature supporting Taylor's position, showing that people distinguish between "response-dependent" evaluations (Taylor's weak evaluations), merely conventional evaluations, and moral evaluations (Taylor's strong evaluations), the last of which are seen even by young children as having an objective quality and being universally applicable.[13]

According to Taylor, the distinguishing feature of a strong evaluation is that it is based on an ontological claim—strong evaluations are based on claims about how the world really is, which is what gives them their power. Evidence from cognitive science and psychology has, however, called into question the directionality of the causal link between ontological belief and moral emotion, suggesting that in many cases the causality may derive from the emotion, with the ontological belief tacked on as an ex post facto justification. Jonathan Haidt et al. (1993) found that when people are presented with verbal scenarios, their affective reactions to them were better predictors of their moral judgments than their claims about harmful consequences, and that people who have a strong negative affective reaction to a scenario often have to struggle to provide a rational justification, with sometimes rather silly results. Similarly, Shaun Nichols (2002) has shown that affectively charged but conventionally neutral scenarios are judged along the same dimensions as moral violations, while Kari Edwards and William von Hippel (1995) found that social attitudes are most effectively changed by altering people's affectively charged intuitions than by rational argument, and that affectively based opinions were held with much more confidence than rationally based ones. In a recent study, Thalia Wheatley and Jonathan Haidt (2005) demonstrated that judgments of both how disgusting and how morally wrong a given behavior is were made more severe by a flash of hypnotically induced queasiness—directly implicating the importance of "gut reactions" for moral judgments.

In defense of his "social intuitionist" approach to moral judgment, Haidt sums up a vast body of literature suggesting that conscious moral reasoning "is usually a post hoc construction, generated after a judgment has been reached" (Haidt 2001, p. 814).[14] These results strongly support the idea that Taylorian strong evaluations arise from affectively laden sensory-motor intuitions, such as disgust, which are then justified through the invocation of ontological claims or rational justification. As Haidt remarks, "Faced with a social demand for a verbal justification, one becomes a lawyer trying to build a case rather than a judge searching for the truth" (p. 814). This has led the neuroscientist Joshua Greene, who has studied the various brain regions involved in moral decision making, to conclude that deontological moral principles are ultimately "a kind of moral confabulation" (Greene 2007). This echoes Nietzsche's famous condemnation of Kant: "Gradually it has become clear to me what every great philosophy so far has been: namely, the personal confession of its author and a kind of involuntary and unconscious memoir; also that the moral (or immoral)

(ID CS4 6 September 2010 12:49 PM) UHP (7×10") Optima J-2344 PEW 61:1 (idp) PMU:(idp[KN]/A)6/9/2010 pp. 80–109 2344_61-1_04 (p. 90)

intentions in every philosophy constituted the real germ of life from which the whole plant had grown" (1886/1966, p. 13).

Now, this position is probably too extreme in least a couple of ways. First of all, it is clear that the causality can go in the opposite direction: explicit beliefs can engender affective responses in otherwise affectively neutral situations. For instance, being touched by a low-caste Indian on the street is not likely to inspire much of an emotional-moral response in me, while it is likely to inspire physical revulsion and moral outrage in a conservative high-caste Hindu, for whom the caste system is part of the normative structure of the cosmos. Even here, though, we should note that this type of culturally specific ontological belief can elicit strong evaluations from an individual only because it is metaphorically tied to basic physiological responses to uncleanliness, contamination, and disgust: physical "mixing" of castes is revolting to a conservative Hindu because of a worldview where lower castes are understood as metaphorically "unclean," which then inspires innate affective responses to pollution and contamination.

Another problem with taking too strongly the affective determination of conscious moral reasoning is that the process can be blocked: it is clear that top-down control, based on rational beliefs, can override affective reactions. This leads us to our next topic.

Rational Top-Down Control is Difficult, Time- and Resource-Consuming, and is Not the Norm

In a cross-cultural study of disgust and moral reactions to various scenarios performed by Haidt et al. (1993), interesting differences were found in the reactions of individuals from cultures varying in their degree of "Westernization" and from various socioeconomic classes. The results suggest that conscious reference to an "ethics of autonomy"—the framework of beliefs concerning the importance of diversity, individual freedom, and rights, et cetera that form the basis of the Western liberal conception of the self—allowed individuals from high socioeconomic classes and living in Westernized cultures to resist converting their affectively negative responses to moralized ones. When presented with certain scenarios—such as a person cutting up the national flag to use as a toilet-cleaning rag, or eating a family dog that had been killed in an accident—these individuals felt at least mild disgust or disapproval, but often overrode this feeling to declare that these actions were not morally wrong, no matter how much they might personally feel uncomfortable with them.[15]

Joshua Greene et al. (2001) document a similar phenomenon in an fMRI study of moral reasoning, where subjects have quite divergent reactions to two versions of a classic thought experiment. In the "trolley" version, the subject is asked to imagine a runaway trolley rolling down a track toward a Y-shaped rail junction, with one person tied to one set of tracks and five people tied to another set. The subject controls a switch that determines which set of tracks the trolley will be diverted to, and most subjects fairly quickly conclude that they would switch the trolley away from the five people and toward the lone person. In the "footbridge" version, the subject is asked to imagine being on a footbridge over a single track to which five people have been

(Id CS4 6 September 2010 12:49 PM) UHP (7×10") Optima J-2344 PEW 61:1 (idp) PMU:(idp[KN]/A)6/9/2010 pp. 80–109 2344_61-1_04 (p. 91)

tied, standing next to a single, but rather large individual. The only way to stop the trolley is to push this person off the bridge to certain death (the subject him- or herself is too small), but this will have the effect of saving the other five. The utilitarian calculus is the same, but in this scenario most subjects say they would not be willing to push the person off the bridge. Greene et al. found that emotional regions of the brain were strongly implicated in subjects' reactions to the footbridge scenario, but not to the trolley version, which suggests that the differing responses are the result of a visceral, negative response to the idea of actively pushing a person off a bridge.

Interestingly, a subset of subjects in this experiment did endorse the "proper" utilitarian response to the footbridge scenario, and the timing of the activation of various brain regions as documented in the fMRI results suggest that this response involved activating brain regions associated with abstract reasoning and cognitive control in order to override the affective responses triggered by the scenario. These sorts of results provide empirical support for the widely held folk belief, endorsed by philosophers since the time of Plato and Xunzi, that the rational faculties *can* supervise and—when appropriate—override the reactions of our more emotional faculties.[16]

An important question for someone interested in an empirically grounded and practically plausible model of morality is, however, the proper degree of *salience* to be given to this type of conscious, rational control. It is clearly possible. What is less clear is how much of an effect it has and can be expected to have on the quotidian functioning of a real moral agent making her way through the world. It is important to note that the sort of conscious overriding of automatic emotional responses documented in the 2001 Greene et al. study significantly interfered with the reaction time of participants, which suggests that cognitive control is a fairly costly and time-consuming process. This accords with the work of Baumeister et al., cited above, documenting the so-called "ego-depletion" phenomenon: conscious supervision is a limited cognitive resource. It is even more important to note that the sort of moral dilemmas that are the staple of deontological and utilitarian theorizing are radically simplified decision-making frames—ethical decision making in the real world takes place in an environment characterized by time pressure, limited and often inaccurate information, indistinct physical and temporal boundaries, and often only limited or entirely nonexistent conscious involvement.[17]

So, part of the problem with the scenarios such as the footbridge case examined by Greene et al. is that they are artificially simplified (in real life, would a person *really* know for sure that the large man's body will stop the train, or that she would be able to get him over the railing in time?); unmotivated (some subjects *say* they would push the man off, but that does not mean they would actually do it in real life, faced by the real human being); and with attention focused on dramatic moments of conscious choice, when in fact most of what counts as moral judgment in real life is probably automatic, or at best semi-conscious. Am I performing a conscious utilitarian calculation (or deontic reasoning) every time I pass a homeless person on my way to the bus? I am a professor and catch a student plagiarizing. The rules say he should get an "F" for the class, but he is the first generation in his family to go to college; my

(ID CS4 6 September 2010 12:49 PM) UHP (7×10") Optima J-2344 PEW 61:1 (idp) PMU:(idp[KN]/A)6/9/2010 pp. 80–109 2344_61-1_04 (p. 92)

"impression" of him is that he is a basically honest, extremely hard-working kid driven to cheat as an act of desperation under extreme pressure; he seems to be experiencing genuine, excruciating remorse; throwing the book at him will result in his scholarship being taken away and expulsion from school; et cetera. When I decide what to do about his case, can I really be said to be performing a utilitarian calculation? What are the numerical parameters I am supposed to be working with? (How many points for remorse? Is it a sliding scale, depending on intensity?)

It seems that these sorts of situations, rather than the trolley or footbridge situations, are what people face most often as the noticeable tip of the iceberg of moral decision making. Then, of course, there is a huge hidden iceberg of moral "choice" that is operating unconsciously all the time when we simply move through the world: are we nice to the check-out girl at the supermarket? Do we notice the old person getting on the bus and give them our seat? Do we pay attention when our spouse is talking to us about something important to her, although we have not yet had our coffee and are thinking intensely about utilitarianism and really just want to get back to our writing? As William Casebeer has noted, the problem with "experimental regimens that isolate 'dry' thinking-about-things-moral from 'wet' here-I-am-doing-moral-things" is that they "can unnecessarily restrict the scope of the neural mechanisms that are activated" (Casebeer 2003) and produce a distorted picture of what real-life moral reasoning is like.[18]

Moral Objectivity Can Be Based on Human Nature, Rather than A Priori *Rational Principles*

One of the more powerful motives for holding on to a rationalist, objectivist model of moral reasoning is the fear of moral relativism. Armed with a substantive picture of human nature, however, we can subscribe to a version of moral intuitionism without having to give up objectivity—at least an anthropocentric objectivity. On the cognitive-evaluative view of emotion, emotions tell us something important about the world, and the basic set of normativity-bestowing emotions experienced by human beings (and many higher mammals) seem to be universal cross-culturally. This is what one would expect from evolution. There is a large literature on the universality of basic human emotions,[19] but I would like to focus here on two moral emotions in particular: empathy and righteous indignation.

Empathy. In recent years, quite a bit has been learned about the neurological mechanisms of empathy in humans and other primates, which appears to involve "mirror neuron" systems and sensory-motor simulation. For primates at least, understanding of another's action occurs when "its observation causes the motor system of the observer to 'resonate'"; in other words, "we understand an action because the motor representation of that action is activated in our brain" (Rizzolatti et al. 2001, p. 661). This sort of understanding-through-enacting can occur even in the absence of actual visual stimuli—it occurs, for instance, when we imagine actions or understand visually obscured but inferred action (Umiltà et al. 2001). Since one is understanding the actions of others by, in a sense, enacting them oneself, the sensory-motor "resonance" that is thus set up can also result in the transmission of

(ID CS4 6 September 2010 12:49 PM) UHP (7×10") Optima J-2344 PEW 61:1 (idp) PMU:(idp[KN]/A)6/9/2010 pp. 80–109 2344_61-1_04 (p. 93)

the behavior itself or its associated emotions. "Contagious yawning" is one example familiar to all of us (Platek et al. 2005), but more subtle types of "emotional contagion" are well documented in primates and humans. In a study by Ulf Dimberg et al. (2000), it was found that exposure to subliminal facial expressions caused a distinct facial electromyographic (EMG) response in the emotion-relevant facial muscles of subjects, and other studies have shown that patients who are congenitally incapable of moving their facial muscles seem to have difficulty understanding emotions conveyed in the expressions of others. This suggests that we unconsciously perceive facial emotions by simulating them ourselves. Coupled with evidence that the sensory-motor simulation of a given emotion can induce the emotion itself (Ekman 2003), this provides a mechanism for explaining the social contagion of moods and feelings.

The resonance induced by action-perception often leads to empathetic feelings and altruistic behavior. A huge body of experiments and behavioral observations with mammals ranging from rats to humans indicates that "individuals of many species are distressed by the distress of a conspecific and will act to terminate the object's distress, even incurring risk to themselves," as Stephanie Preston and Frans de Waal explain in their review of the literature (Preston and de Waal 2002, p. 1). They note that this altruistic drive shows a familiarity effect (altruism is more likely when the object is a known individual), is sensitive to past experience (i.e., whether or not the animal has suffered from similar trouble), and is somewhat dependent on "cue salience" (the strength of the perceptual signal). This all suggests that "attended perception [of the object's suffering] automatically activates the subject's representations of the state, situation, and object, and that activation of these representations automatically primes or generates the associated autonomic and somatic responses, unless inhibited" (p. 4). There is considerable evidence that human psychopathy involves a breakdown of this imagery-based empathy-altruism system, with clear deficits in mirror-neuron and emotional regions of the brain (Blair 2001).

Righteous indignation ("Justice"). Many studies have found prosocial behavior among humans even in anonymous, non-iterated games (Gintis et al. 2003). This suggests that there exists, in at least a subset of the human population, a general bias toward cooperation. Experimental manipulation of the parameters of economic games has shown, however, that cooperation is usually best maintained when players have a way to monitor or predict the level of cooperation of others, and therefore be in a position to punish "free-riders" who fail to cooperate (see, e.g., Fehr and Gächter 2000). Indeed, theoretical models suggest that cooperation between non-kin could only evolve in the presence of mechanisms to prevent free-riding—otherwise free-riders would inevitably profit at the expense of cooperators and drive cooperators out of the gene pool.[20]

One of these anti-free-rider mechanisms seems to be a specialized "cheater detection" module in the brain—a subsystem dedicated to identifying violations of the social contract—that seems to be sensitive to intentionality and accrual of benefit, be present cross-culturally, and be selectively impaired by brain damage.[21] This "cheater

(ID CS4 6 September 2010 12:49 PM) UHP (7×10") Optima J-2344 PEW 61:1 (idp) PMU:(idp[KN]/A)6/9/2010 pp. 80–109 2344_61-1_04 (p. 94)

detection" module appears to be paired with a strong "cheater punishment" sentiment that motivates individuals to punish cheaters even at significant cost to themselves (so-called "altruistic punishment").[22] Studies of the behavior of subjects in economic simulations such as the "Ultimatum Game" show that they are easily provoked into rationally unjustifiable punishing behaviors when confronted with apparent injustice. In a typical Ultimatum Game scenario, subject 1 is given a sum of money and told to split it with subject 2 in any proportion he or she desires, with the sole proviso that subject 2 can reject the offer, in which case neither subject gets any money. According to classical economic theory there is only one rational strategy: given one hundred dollars to split, subject 1 should offer $0.01 to subject 2, keeping $99.99 for him/herself, and subject 2 should accept the offer, since both subjects end up better off financially than before. Not all people will react this way: subjects on the receiving end of the Ultimatum Game indignantly reject offers that are viewed as grossly unfair, with the usual rejection threshold hovering around 20 to 30 percent.[23]

A fascinating fMRI Ultimatum Game study by Alan Sanfrey et al. (2003) showed that low offers made by human partners were rejected at a significantly higher rate than low offers from a computer (which suggests that this sentiment is particularly sensitive to unfair treatment from other humans), and that the rejection or acceptance by subjects of unfair offers hinged on the relative activation of the anterior insula (associated with negative emotion states such as anger and disgust) versus the dorsolateral prefrontal cortex and anterior cingulated cortex (associated with abstract reasoning and top-down cognitive control). In other words, subjects had a clear negative emotional response to unfair offers (especially from other humans), but in some subjects this reaction was suppressed by more "rational" brain systems in pursuit of the general goal of money accumulation. The behavior of subjects in scenarios such as the Ultimatum Game is echoed by the behavior of real and simulated juries studied by Daniel Kahneman and his colleagues in their studies of "retributive punishment": they found an almost perfect correlation between degree of emotional outrage and mean punishment levied on real or imagined law violators, regardless of the actual damage these violations had caused (Kahneman et al. 1998).

This desire to punish norm violators even at cost to oneself is clearly driven by powerful positive emotions: a study by Dominique De Quervain et al. (2004) showed that people derive hedonistic satisfaction from punishing norm violators in situations like the Ultimatum Game. This sentiment also seems to be distinct from other altruistic sentiments, having a more absolutist and less context-sensitive quality. Rick O'Gorman et al. (2005) found that while the desire to altruistically *help* was very sensitive to information regarding genetic relatedness or potential future interactions, the desire to altruistically *punish* was entirely insensitive to such information—indicating "an important psychological difference between altruistic punishment and altruistic helping." A recent study by Björn Wallace et al. (2007) argues that, like other personality traits such as introversion or conservatism, patterns of response in the Ultimatum Game have a genetic component, strongly suggesting that they are part of evolved human cognitive architecture. Finally, like empathy, this emotion of

(Id CS4 6 September 2010 12:49 PM) UHP (7×10") Optima J-2344 PEW 61:1 (idp) PMU:(idp[KN]/A)6/9/2010 pp. 80–109 2344_61-1_04 (p. 95)

righteous indignation may not be confined to humans. Studies performed by Sarah Brosnan and Frans de Waal (2003) indicate that something like a sense of injustice is present in capuchin monkeys, who are willing to incur costs to themselves in order to express their displeasure with unfair food-distribution reward systems.[24]

So, it appears that even if we abandon the attempt to find an objective, *a priori* rational grounding for our moral intuitions, we are not left floating in a relativist vacuum. This fact has been appreciated by philosophers such as Martha Nussbaum, who has argued that we can rely on a set of pan-human "grounding experiences" (and the emotions that go along with them) as a basis for objectivity (Nussbaum 1988). Of course, one of the peculiar features of moral intuitions is that they present themselves to us as objective in an *absolute* sense, and scholars taking an evolutionary approach to ethics have argued that this "illusion" of moral realism is motivationally necessary (and thus built in by evolution) in order to enable complex social interactions (Ruse and Wilson 1986). How our deepest moral intuitions can be viewed intellectually as anthropocentric illusions and yet still serve their motivational function is an interesting psychological question, but this sort of dual consciousness does not appear to be impossible in practice. I would tend to agree with a modified form of Charles Taylor's claim that human-centered concepts that cannot be dispensed with—in other words, that we cannot help but make reference to in moving through and making sense of our world—are as "real" as is reasonable to hope for (Taylor 1989, p. 56).

Empirically Responsible Philosophy, Virtue Ethics, and Some Mencian Parallels

Allow me to quote an apropos call-to-arms issued by David Hume over two hundred years ago:

> Men are now cured of their passion for hypotheses and systems in natural philosophy, and will hearken to no arguments but those which are derived from experience. It is full time that they should attempt a like reformation in all moral disquisitions; and reject every system of ethics, however subtle or ingenious, which is not founded on fact and observation. (Hume 1777/1976, pp. 174–175)

What I would like to suggest is that the body of empirical evidence emerging from cognitive science, cognitive linguistics, neuroscience, social psychology, and primatology that I have reviewed above suggests that the so-called "virtue ethical" model best describes how real human beings actually *engage* in moral reasoning (a descriptive claim), and therefore provides us with the best framework for formulating a psychologically realistic model of moral reasoning and moral education (a normative claim).[25] Since we are invoking Hume, we must, of course, guard against the error of slipping from "is" into "ought." I think, however, that psychological feasibility is an important desideratum for any ethical theory. If deontology and utilitarianism require us to think or behave in manners that are simply not possible or sustainable in quotidian life, this should temper our enthusiasm for adopting them as moral ideals.

(ID CS4 6 September 2010 12:49 PM) UHP (7×10") Optima J-2344 PEW 61:1 (idp) PMU:(idp[KN]/A)6/9/2010 pp. 80–109 2344_61-1_04 (p. 96)

Of course, the modern revival of virtue ethics has traditionally looked back to Aristotle for inspiration, and many of the features of human psychology discussed above can easily be accommodated by the Aristotelian conception of the self. As many scholars of Chinese thought have argued, however,[26] there are other virtue ethicists besides Aristotle, and looking beyond Aristotle to traditions such as early Confucianism can give us a broader, richer picture of what a virtue ethic in action might look like. Because this essay is intended for an audience familiar with Chinese thought and unfamiliar with modern cognitive science, I have focused most of my attention on exploring the latter topic; my hope is that the previous outline of the state of field in cognitive science has made immediately apparent to my colleagues in early Chinese thought—and those concerned with virtue ethics in general—what cognitive science has to offer us. Here I would like to conclude by very briefly connecting at least a few of the dots, suggesting some of the ways in which one early Confucian model of moral reasoning and moral education, that of Mencius, anticipates and extends the observations about human psycho-physiology described above, and thus might serve as a resource for formulating a modern, empirically responsible ethical system.[27]

The most obvious point of congruence is the foundational role that emotions play in Mencian moral reasoning and decision making. The Mencian "sprouts" (*duan* 端) are feelings or "movements of the heart" (*xin* 心) that represent affective, yet intelligent, responses to the perception of value in the world. Even the specific identification of these sprouts by Mencius seems to be a good first approximation of a plausible catalog of innate human moral emotions. The feeling of "not being able to bear it" (*burenzhixin* 不忍之心)—the sprout of the virtue of benevolence or empathy (*ren* 仁)—is clearly, as discussed above, a basic mammalian moral emotion.[28] The imagined indignant refusal of a life-saving bowl of soup offered in an insulting manner is taken by Mencius in 6A:10 to reveal the existence of an innate refusal to accept practically beneficial but unrighteous behavior, which in turn is the sprout of the virtue of rightness (*yi* 義). This is not terribly unlike the documented responses of humans, and perhaps other primates, to insultingly low offers in the Ultimatum Game, although, of course, giving up one's life is several magnitudes more serious a test of one's righteousness than forgoing the prospect of a modest sum of money. Similarly, the "heart of deference" (2A:6) that Mencius sees as the sprout of ritual propriety (*li* 禮) might be compared to the foundational role of disgust reactions in moral judgments, as documented by Paul Rozin, Jonathan Haidt, and others.

Mencian sprouts are merely moral potentialities, of course, and Mencian moral education focuses on the gradual strengthening and "extension" (*tui* 推) of these incipient virtues.[29] Imagination is clearly central to this process, whether in recalling and dwelling upon one's prior somatic-emotional states, as when Mencius asks the king in 1A:7 to recall his sparing of the ox, or the summoning up of purely imaginary scenarios, such as the famous child and the well from 2A:6. Once the proper feelings have been imagined or imaginatively recreated, the next step in the Mencian

(ID CS4 6 September 2010 12:49 PM) UHP (7×10") Optima J-2344 PEW 61:1 (idp) PMU:(idp[KN]/A)6/9/2010 pp. 80–109 2344_61-1_04 (p. 97)

process is to extend them through a process of sympathetic projection, guided by metaphor and analogy:

> Treat the aged of your own family in a manner that respects their seniority, and then cause this treatment to reach the aged of other families. Treat the young ones of your family in a manner appropriate to their youth, and then cause this treatment to reach the young of other families. Once you are able to do this, you will have the world in the palm of your hand. . . . All that is required is to pick up this heart here and apply it to what is over there. (1A : 7)

Although many scholars have portrayed this process of Mencian extension as a rational equation of logically similar situations,[30] it seems more accurate to understand it as a process of "analogical resonance," involving "emotional resonance not cognitive similarity" (Ivanhoe 2002, p. 226).[31] In addition to his skill as a moral psychoanalyst, Mencius has at his disposal the standard tools of Confucian moral self-cultivation—ritual, music, the inspiring examples of the sage kings—which clearly involve a kind of analog, somatic-emotional prototype modeling.[32] Moreover, it is precisely these sorts of cultural templates for thought and behavior that could be expected to ensure proper behavior in creatures guided by habit and automaticity: the absence of an all-powerful, all-knowing cognitive commander-in-chief matters little if ritual and custom are there to catch you.

The importance of prototype and analogical reasoning extends to Mencius' philosophical struggles with his opponents. Debates such as those between Mencius and Gaozi in 6A : 1–4, for instance, are best understood as battles of competing metaphors, each bearing with it specific normative values,[33] rather than an "atrociously inept and unconvincing" (Hansen 2000, p. 188) attempt at abstract, rational argumentation. If human nature really is like a swirling pool, and moral education like the directing of this water in a useful direction, who are we to quibble with Gaozi? Mencius wins us over, not because, in some objective way, he "truly understood water" (Allen 1997, p. 42)—whatever that could even mean—but because the editors of the text have structured the debate so as to leave us feeling that Mencius' metaphorical framing is the most appropriate. The content of Mencius' critique of Mohism also anticipates the argument that any viable ethic needs to be a psychologically realistic one: asking people to be better than they are capable of being is to ask for trouble.

Finally, Mencius' argument that human beings share a common, innate "taste" for morality (6A : 7) resonates with the idea that morality can be grounded in empirical claims about human nature. Mencius' particular picture of this shared taste may be overly rosy: a more extensive, well-rounded account of innate human emotions than the one presented above would have to include our propensity for horrific out-group violence; our powerful desires for wealth, power, comfort, and sex; and the prevalence of opportunistic selfishness and convenient self-deception.[34] Nonetheless, the idea that morality involves the selective strengthening and extending of urges arising from ourselves, rather than being "welded on from the outside" (6A : 6), can serve as a starting point for a morality grounded not in Kant's "starry Heaven above," but in the messy, contingent, and yet inescapable facticity of our embodiment.

(ID CS4 6 September 2010 12:49 PM) UHP (7×10") Optima J-2344 PEW 61:1 (idp) PMU:(idp[KN]/A)6/9/2010 pp. 80–109 2344_61-1_04 (p. 98)

Notes

This article has benefited greatly from helpful comments provided by the anonymous referees, as well as audience members at the International Conference on Chinese Philosophy in Analytical Perspectives, National Chengchi University, Taipei, Taiwan, September 2005, and the Conference on Chinese Philosophy and Moral Psychology, Hong Kong University of Science and Technology, December 2007. Thanks also to Owen Flanagan, Jon Haidt, Josh Greene, Jess Tracy, and Jonathan Schooler for helping me to grasp the relevant cognitive science and social psychology and to work out some of the issues that I address here. This research was undertaken, in part, thanks to funding from the Canada Research Chairs Program.

1 – For a recent and very lucid summary of relationships between the various species of contemporary ethical theories, and their relationship to early Chinese thought, see Van Norden 2007.

2 – Two anonymous reviewers of this article (as well as several audience members when I have presented versions of this material at conferences) have helpfully—and correctly—pointed out that more recent proponents of both deontology and utilitarianism have acknowledged an important role for intuitive, implicit cognitive processes, sometimes confining explicit, algorithmic reasoning to a critical meta-level, which may only be invoked when conflicts arise or justifications need to be provided. While certainly more psychologically realistic, this still begs the question of how the behavioral *desiderata* arrived at through deontological or utilitarian reasoning are to be built into automatic, everyday cognition, which is an issue that virtue ethics uniquely and explicitly addresses. For more on this I refer the reader to my characterization of Confucian virtue ethics as "time-delayed cognitive control," in two works, Slingerland under review and Slingerland in preparation.

3 – For a more detailed account, see Slingerland 2008, chaps. 1 and 3, as well as Gibbs 2006.

4 – Also see Johnson 1987, Lakoff 1987, and Langacker 1987 for similar arguments that linguistic representations have an analog, spatial component, as well as Damasio's claim that "images are the main content of our thought" (Damasio 1994, p. 107).

5 – For reviews see Warrington and Shallice 1984; Barsalou 1999, pp. 579–580; Martin and Chao 2001; and the essays in Pecher and Zwaan 2005.

6 – Rosch 1973 and Rosch et al. 1976; also see Lakoff 1987.

7 – For just a sampling, see Damasio 1994 and 2003, De Sousa 1987, Ekman and Davidson 1994, Lazarus 1991, Le Doux 1996, Nussbaum 2001, and Solomon 2003 and 2004.

8 – See, for example, Kahneman, Slovic, and Tversky 1982 and Kahneman and Tversky 2000. Also see Gigerenzer and Selten 2001 for an introduction to the

(ID CS4 6 September 2010 12:49 PM) UHP (7×10") Optima J-2344 PEW 61:1 (idp) PMU:(idp[KN]/A)6/9/2010 pp. 80–109 2344_61-1_04 (p. 99)

argument that "fast and frugal" heuristics often outperform general-purpose, time-consuming, and "information-greedy" optimizing strategies, especially in the specific situations of partial knowledge and computational limitations for which they are designed.

9 – For a recent literature survey and account of the role and power of the "adaptive unconscious," see Timothy Wilson 2002.

10 – See Johnson 1993, pp. 91–98, for a discussion of this work.

11 – On the importance of metaphoric framing in private reasoning and public debate, see especially Lakoff 1996, Chilton 1996, Coulson 2001, Fauconnier and Turner 2002, Beer and De Landtsheer 2004, and Slingerland et al. 2007.

12 – Cf. Lakoff and Johnson's argument that normative judgments arise from background feelings, and that bodily ease or "well-being" is the basic source domain for our understanding of "goodness" (Lakoff and Johnson 1999, pp. 290–292). For representative recent works on morality and emotion, see Sober and Wilson 1998, the essays collected in Solomon 2004, Nichols 2004, and Prinz 2005 and 2006.

13 – See the literature reviewed in Nichols 2002, pp. 221–222, and Nichols and Folds-Bennett 2003 for results with four- to six-year-old children.

14 – See Haidt 2007 for a more recent literature review.

15 – See Jones 2007 for a recent review article on disgust and moral judgments.

16 – On this topic, also see Greene et al. 2004.

17 – Daniel Dennett's example of the philosophy fellowship competition is a helpful illustration of some of the features of "myopic and time-pressured" real-time decision making (Dennett 1995, pp. 495–502).

18 – He and Patricia Churchland have argued for a more "ecologically valid experimental regime" that takes into account that real moral reasoning is "hot" (affective states are a crucial part), social (decisions not made in a social vacuum, subject to social cues), distributed (embedded in a large web of stimulation), organic (context sensitive), genuine (personally involved rather than abstract), and directed (about actual things in the world) (Casebeer and Churchland 2003, pp. 187–188).

19 – See especially the work of Paul Ekman (Ekman 1982 and 2003, Ekman and Davidson 1994).

20 – See Henrich and Boyd 2001 for a discussion and literature review.

21 – See Cosmides and Tooby 2005, and Sugiyama, Tooby, and Cosmides 2002.

22 – Price, Cosmides, and Tooby 2002; similar results were found with hunter-horticulturalists in the Ecuadorian Amazon, in Price 2005.

(ID CS4 6 September 2010 12:49 PM) UHP (7×10") Optima J-2344 PEW 61:1 (idp) PMU:(idp[KN]/A)6/9/2010 pp. 80–109 2344_61-1_04 (p. 100)

23 – See Henrich et al. 2006 for a study of subjects from five continents, representing the full range of human modes of production, that suggests the presence of a widespread tendency to punish unfair offers that is then calibrated culturally, with Ultimatum Game offer rejection thresholds varying significantly from culture to culture.

24 – See, however, a recent study by Jensen et al. (2007) suggesting that chimpanzees seem to lack this response, and in fact perform like classic rational actors in a version of the Ultimatum Game.

25 – Observations along these lines have been made by neuroscientifically literate philosophers such as Flanagan (1991), Churchland (1998), and Casebeer (2003). John Doris (1998 and 2002) and Gilbert Harman (1999) have famously argued that findings in social psychology suggest that stable human character traits do not exist, which of course would call into question the very *raison d'être* of virtue ethics. In fact, large-scale meta-analyses—for example, Roberts et al. 2007—show that while situational effects can be quite strong, stable personality traits have at least as strong an effect. Also see Kupperman 2001, Kamtekar 2004, Hutton 2006, and Slingerland (under review) for relevant responses to Doris' and Harman's positions.

26 – Stephen Wilson 1995, Ivanhoe 2000, Slingerland 2001, and Van Norden 2007.

27 – For a more fully developed exploration of this theme, the reader is referred to Munro 2005 and Slingerland (in preparation).

28 – The fact that empathy is often only felt in quite narrow contexts, or is insufficiently translated into actual action, is a basic concern of early Confucian moral philosophy, captured as well in Confucius' emphasis upon the importance of *shu* 恕 or "sympathetic understanding."

29 – This account of Mencian self-cultivation is derived from Ivanhoe 2000, to which the reader is referred for a more complete account.

30 – See, for example, Eric Hutton's claim that Mencius' "basic method is to make a certain claim P about some entity A or set of entities {A, B, C, . . .} and then ask whether or not the same claim applies to some other entity X" (Hutton 2002, p. 169).

31 – Cf. Wong 2002 and Van Norden 2007.

32 – Cf. Robert Eno's comment regarding the dependence of Mencian "character ethics" on "a methodology of *verstehen,* an emphatic grasp of virtue perspectives cultivated through hermeneutic probing of historical narratives" (Eno 2002, p. 190).

33 – See Slingerland 2005 for an in-depth defense of this position.

34 – I thank Owen Flanagan (personal communication) for this point.

(ID CS4 6 September 2010 12:49 PM) UHP (7×10") Optima J-2344 PEW 61:1 (idp) PMU:(idp[KN]/A)6/9/2010 pp. 80–109 2344_61-1_04 (p. 101)

References

Allen, Sarah. 1997. *The Way of Water and Sprouts of Virtue.* Albany: State University of New York Press.

Bargh, John, and Tanya Chartrand. 1999. "The Unbearable Automaticity of Being." *American Psychologist* 54 : 462–479.

Barsalou, Laurence. 1999. "Perceptual Symbol Systems." *Behavioral and Brain Sciences* 22 : 577–609.

Baumeister, Roy. 1984. "Choking Under Pressure: Self-consciousness and Paradoxical Effects of Incentives on Skillful Performance." *Journal of Personality and Social Psychology* 46 : 610–620.

Baumeister, Roy, Ellen Bratslavsky, Mark Muraven, and Dianne Tice. 1998. "Ego Depletion: Is the Active Self a Limited Resource?" *Journal of Personality and Social Psychology* 74 : 1252–1265.

Beer, Francis, and Christ'l de Landtsheer, eds. 2004. *Metaphorical World Politics.* East Lansing: Michigan State University Press.

Blair, R. James. 2001. "Neurocognitive Models of Aggression, the Antisocial Personality Disorders, and Psychopathy." *Journal of Neurology, Neurosurgery, and Psychiatry* 71 : 727–731.

Brosnan, Sarah, and Frans de Waal. 2003. "Monkeys Reject Unequal Pay." *Nature* 425 : 297–299.

Casebeer, William D. 2003. "Moral Cognition and Its Neural Constituents." *Nature Reviews Neuroscience* 4.

Casebeer, William, and Patricia Churchland. 2003. "The Neural Mechanisms of Moral Cognition: A Multi-aspect Approach to Moral Judgment and Decision-making." *Biology and Philosophy* 18 : 169–194.

Chilton, Paul. 1996. Security Metaphors: Cold War Discourse from Containment to Common House. New York: Peter Lang.

Churchland, Paul. 1998. "Toward a Cognitive Neurobiology of the Moral Virtues." *Topoi* 17 : 83–96.

Coleman, Linda, and Paul Kay. 1981. "Prototype Semantics: The English Word Lie." *Language* 57 : 26–44.

Cosmides, Leda, and John Tooby. 2005. "Neurocognitive Adaptations Designed for Social Exchange." In David Buss, ed., *The Handbook of Evolutionary Psychology,* pp. 584–627. Hoboken, NJ: John Wiley and Sons.

Coulson, Seana. 2001. *Semantic Leaps.* Cambridge: Cambridge University Press.

Damasio, Antonio. 1989. "The Brain Binds Entities and Events by Multiregional Activation from Convergence Zones." *Neural Computation* 1 : 123–132.

(ID CS4 6 September 2010 12:49 PM) UHP (7×10") Optima J-2344 PEW 61:1 (idp) PMU:(idp[KN]/A)6/9/2010 pp. 80–109 2344_61-1_04 (p. 102)

———. 1994. Descartes' Error: Emotion, Reason, and the Human Brain. New York: G. P. Putnam's Sons.

———. 2003. In Search of Spinoza: Joy, Sorrow, and the Feeling Brain. New York: Harvest.

Doris, John. 1998. "Persons, Situations, and Virtue Ethics." *Noûs* 32 : 504–530.

———. 2002. *Lack of Character: Personality and Moral Behavior.* New York: Cambridge University Press.

De Sousa, Ronald. 1987. *The Rationality of Emotion.* Cambridge University Press.

De Quervain, Dominique, Urs Fischbacher, Valerie Treyer, Melanie Schellhammer, Ulrich Schnyder, Alfred Buck, and Ernst Fehr. 2004. "The Neural Basis of Altruistic Punishment." *Science* 305.5688 : 1254–1258.

Dennett, Daniel. 1995. Darwin's Dangerous Idea: Evolution and the Meaning of Life. New York: Simon and Schuster.

Dimberg, Ulf, Monika Thunberg, and Kurt Elmehed. 2000. "Unconscious Facial Reactions to Emotional Facial Expressions." *Psychological Science* 11 : 86–89.

Edwards, Kari, and William von Hippel. 1995. "Hearts and Minds: The Priority of Affective Versus Cognitive Factors in Person Perception." *Personality and Social Psychology Bulletin* 21 : 996–1011.

Ekman, Paul. 1982. *Emotion in the Human Face.* 2nd ed. Cambridge: Cambridge University Press.

———. 2003. Emotions Revealed: Recognizing Faces and Feelings to Improve Communication and Emotional Life. Times Books.

Ekman, Paul, and Richard Davidson. 1994. *The Nature of Emotion.* New York: Oxford University Press.

Eno, Robert. 2002. "Casuistry and Character in the *Mencius.*" In Alan Chan, ed., *Mencius: Contexts and Interpretations,* pp. 189–215. Honolulu: University of Hawai'i Press.

Fauconnier, Gilles, and Mark Turner. 2002. The Way We Think: Conceptual Blending and the Mind's Hidden Complexities. New York: Basic Books.

Fehr, Ernst, and Simon Gächter. 2000. "Fairness and Retaliation: The Economics of Reciprocity." *Journal of Economic Perspectives* 14 : 159–181.

Flanagan, Owen. 1991. Varieties of Moral Personality: Ethics and Psychological Realism. Cambridge, MA: Harvard University Press.

Gazzaniga, Michael. 1998. *The Mind's Past.* Berkeley: University of California Press.

Gibbs, Raymond. 2006. *Embodiment and Cognitive Science.* Cambridge: Cambridge University Press.

(Id CS4 6 September 2010 12:49 PM) UHP (7×10") Optima J-2344 PEW 61:1 (idp) PMU:(idp[KN]/A)6/9/2010 pp. 80–109 2344_61-1_04 (p. 103)

Gigerenzer, Gerd, and Reinhard Selten, eds. 2001. *Bounded Rationality: The Adaptive Toolbox.* Cambridge, MA: MIT Press.

Gintis, Herbert, Samuel Bowles, Robert Boyd, and Ernst Fehr. 2003. "Explaining altruistic behavior in humans." *Evolution and Human Behavior* 24 : 153–172.

Grady, Joseph, Todd Oakley, and Seana Coulson. 1999. "Blending and Metaphor." In Raymond Gibbs and Gerard Steen, eds., *Metaphor in Cognitive Linguistics,* pp. 101–124. Philadelphia, PA: John Benjamins Publishing Company.

Greene, Joshua. 2007. "The Secret Joke of Kant's Soul." In W. Sinnott-Armstrong, ed., *Moral Psychology,* vol. 3, *The Neuroscience of Morality: Emotion, Disease, and Development,* pp. 35–79. Cambridge, MA: MIT Press.

Greene, Joshua, R. Brian Sommerville, Leigh E. Nystrom, John M. Darley, and Jonathan D. Cohen. 2001. "An fMRI Investigation of Emotional Engagement in Moral Judgment." *Science* 293 : 2105–2108.

Greene, Joshua, Leigh Nystrom, Andrew Engell, John Darley, and Jonathan Cohen. 2004. "The Neural Bases of Cognitive Conflict and Control in Moral Judgment." *Neuron* 44 : 389–400.

Haidt, Jonathan. 2001. "The Emotional Dog and Its Rational Tail: A Social Intuitionist Approach to Moral Judgment." *Psychological Review* 108 : 813–834.

———. 2007. "The New Synthesis in Moral Psychology." *Science* 316 : 998–1002.

Haidt, Jonathan, Silvia Koller, and Maria Dias. 1993. "Affect, Culture, and Morality, or Is It Wrong to Eat Your Dog?" *Journal of Personality and Social Psychology* 65 : 613–628.

Hansen, Chad. 2000. A Daoist Theory of Chinese Thought: A Philosophical Interpretation. 2nd ed. New York: Oxford University Press.

Harman, Gilbert. 1999. "Moral Philosophy Meets Social Psychology: Virtue Ethics and the Fundamental Attribution Error." *Proceedings of the Aristotelian Society* 99 : 315–331.

Henrich, Joseph, and Robert Boyd. 2001. "Why People Punish Defectors: Weak Conformist Transmission Can Stabilize Costly Enforcement of Norms in Cooperative Dilemmas." *Journal of Theoretical Biology* 208 : 79–89.

Henrich, Joseph, Richard McElreath, Abigail Barr, Jean Ensminger, Clark Barrett, Alexander Bolyanatz, Juan Camilo Cardenas, Michael Gurven, Edwins Gwako, Natalie Henrich, Carolyn Lesorogol, Frank Marlowe, David Tracer, and John Ziker. 2006. "Costly Punishment across Human Societies." *Nature* 312 : 1767–1770.

Hume, David. 1777/1976. Enquiries Concerning Human Understanding and Concerning the Principles of Morals. 3d ed. Oxford: Clarendon.

Hutton, Eric. 2002. "Moral Connoisseurship in Mengzi." In Liu and Ivanhoe 2002, pp. 163–186.

(ID CS4 6 September 2010 12:49 PM) UHP (7×10") Optima J-2344 PEW 61:1 (idp) PMU:(idp[KN]/A)6/9/2010 pp. 80–109 2344_61-1_04 (p. 104)

———. 2006. "Character, Situationalism, and Early Confucian Thought." *Philosophical Studies* 127 : 37–58.

Ivanhoe, Philip J. 2000. *Confucian Moral Self-Cultivation.* 2nd ed. Indianapolis and Cambridge: Hackett Publishing Company.

———. 2002. "Confucian Self Cultivation and Mengzi's Notion of Extension." In Liu and Ivanhoe 2002, pp. 221–241.

Jensen, Keith, Josep Call, and Michael Tomasello. 2007. "Chimpanzee are Rational Maximizers in an Ultimatum Game." *Science* 318 : 107–109.

Johnson, Mark. 1987. The Body in the Mind: The Bodily Basis of Meaning, Imagination and Reason. Chicago: University of Chicago Press.

———. 1993. Moral Imagination: Implications of Cognitive Science for Ethics. Chicago: University of Chicago Press.

Jones, Dan. 2007. "The Depths of Disgust." *Nature* 447 : 768–771.

Kahneman, Daniel, David Schkade, and Cass Sunstein. 1998. "Shared Outrage and Erratic Rewards: The Psychology of Punitive Damages." *Journal of Risk and Uncertainty* 16 : 49–86.

Kahneman, Daniel, Paul Slovic, and Amos Tversky, eds. 1982. *Judgement under Uncertainty: Heuristics and Biases.* Cambridge: Cambridge University Press.

Kahneman, Daniel, and Amos Tversky, eds. 2000. *Choices, Values, and Frames.* Cambridge: Cambridge University Press.

Kamtekar, Rachana. 2004. "Situationalism and Virtue Ethics on the Content of Our Character." *Ethics* 114 : 458–491.

Kövecses, Zoltán. 2002. *Metaphor: A Practical Introduction.* Oxford: Oxford University Press.

Kunst-Wilson, William, and Robert Zajonc. 1980. "Affective Discrimination of Stimuli that Cannot Be Recognized." *Science* 207 : 557–558.

Kupperman, Joel. 2001. "The Indispensability of Character." *Philosophy* 76 : 239–250.

Lakoff, George. 1987. Women, Fire and Dangerous Things: What Categories Reveal about the Mind. Chicago: University of Chicago Press.

———. 1996. Moral Politics: What Conservatives Know that Liberals Don't. Chicago: University of Chicago Press.

Lakoff, George, and Mark Johnson. 1999. Philosophy in the Flesh: The Embodied Mind and Its Challenge to Western Thought. New York: Basic Books.

Langacker, Ronald. 1987. *Foundations of Cognitive Grammar.* Vol. 1, *Theoretical Prerequisites.* Stanford: Stanford University Press.

Lazarus, Richard. 1991. *Emotion and Adaptation.* New York: Oxford University Press.

(ID CS4 6 September 2010 12:49 PM) UHP (7×10") Optima J-2344 PEW 61:1 (idp) PMU:(idp[KN]/A)6/9/2010 pp. 80–109 2344_61-1_04 (p. 105)

Le Doux, Joseph. 1996. The Emotional Brain: The Mysterious Underpinnings of Emotional Life. New York: Simon and Schuster.

Liu Xiusheng and P. J. Ivanhoe, eds. 2002. *Essays on the Moral Philosophy of Mengzi.* Indianapolis and Cambridge: Hackett Publishing Company.

Loewenstein, George, Elke Weber, Christopher Hsee, and Ned Welch. 2001. "Risk as Feelings." *Psychological Bulletin* 127 (2): 267–286.

Martin, Alex, and Linda Chao. 2001. "Semantic Memory and the Brain: Structure and Processes." *Current Opinion in Neurobiology* 11 : 194–201.

Mealey, Linda, Christiopher Daood, and Michael Krage. 1996. "Enhanced Memory for Faces of Cheaters." *Ethology and Sociobiology* 17 : 119–128.

Munro, Donald. 2005. A Chinese Ethics for the New Century: The Ch'ien Mu Lectures in History and Culture, and Other Essays on Science and Confucian Ethics. Hong Kong: Chinese University of Hong Kong.

Muraven, Mark, Dianne Tice, and Roy Baumeister. 1998. "Self-Control as a Limited Resource: Regulatory Depletion Patterns." *Journal of Personality and Social Psychology* 74 : 774–789.

Nichols, Shaun. 2002. "Norms with Feeling: Towards a Psychological Account of Moral Judgment." *Cognition* 84 : 221–236.

———. 2004. Sentimental Rules: On the Natural Foundations of Moral Judgment. New York: Oxford University Press.

Nichols, Shaun, and Trisha Folds-Bennett. 2003. "Are Children Moral Objectivists? Children's Judgments about Moral and Response-Dependent Properties." *Cognition* 90 : B23–B32.

Nietzsche, Friedrich. 1886/1966. *Beyond Good and Evil.* Translated by Walter Kaufmann. New York: Vintage.

Nisbett, Richard, and Timothy Wilson. 1977. "Telling More than We Can Know: Verbal Reports on Mental Processes." *Psychological Review* 84 : 231–259.

Nussbaum, Martha. 1988. "Non-relative Virtues: An Aristotelian Account." In *Midwest Studies in Philosophy Volume XIII: Ethical Theory: Character and Virtue,* edited by Peter A. French, pp. 32–53. University of Notre Dame Press.*

———. 1990. Love's Knowledge: Essays on Philosophy and Literature. Oxford: Oxford University Press.

———. 2001. *Upheavals of Thought: The Intelligence of the Emotions.* Cambridge: Cambridge University Press.

O'Gorman, Rick, David Sloan Wilson, and Ralph Miller. 2005. "Altruistic Punishment and Helping Differ in Sensitivity to Relatedness, Friendship, and Future Interactions." *Evolution and Human Behavior* 26 : 375–387.

(ID CS4 6 September 2010 12:49 PM) UHP (7×10") Optima J-2344 PEW 61:1 (idp) PMU:(idp[KN]/A)6/9/2010 pp. 80–109 2344_61-1_04 (p. 106)

Ortony, Andrew, Gerald Clore, and Allan Collins. 1990. *The Cognitive Structure of Emotion.* Cambridge: Cambridge University Press.

Pecher, Diane, and Rolf Zwaan, eds. 2005. Grounding Cognition: The Role of Perception and Action in Memory, Language and Thinking. Cambridge: Cambridge University Press.

Platek, Steven, Feroze Mohamed, and Gordon Gallup Jr. 2005. "Contagious Yawning and the Brain." *Cognitive Brain Research* 23 : 448–453.

Preston, Stephanie, and Frans de Waal. 2002. "Empathy: Its Ultimate and Proximate Bases." *Behavioral and Brain Sciences* 25 : 1–72.

Price, Michael. 2005. "Punitive Sentiment among the Shuar and in Industrialized Societies: Cross-cultural Similarities." *Evolution and Human Behavior* 26 : 279–287.

Price, Michael, Lena Cosmides, and John Tooby. 2002. "Punitive Sentiment as an Anti-Free Rider Psychological Device." *Evolution and Human Behavior* 23 : 203–231.

Prinz, Jesse. 2005. "Passionate Thoughts: The Embodiment of Moral Concepts." In Pecher and Zwaan 2005, pp. 93–114. Cambridge: Cambridge University Press.

———. 2006. *The Emotional Construction of Morals.* New York: Oxford University Press.

Rizzolatti, Giacomo, Leonardo Fogassi, and Vittorio Gallese. 2001. "Neurophysiological Mechanisms Underlying the Understanding and Imitation of Action." *Nature Reviews Neuroscience* 2 : 661–670.

Roberts, Brent, Nathan Kuncel, Rebecca Shiner, Avshalom Caspi, and Lewis Goldberg. 2007. "The Power of Personality: The Comparative Validity of Personality Traits, Socioeconomic Status, and Cognitive Ability for Predicting Important Life Outcomes." *Perspectives on Psychological Science* 2 : 313–345.

Rosch, Eleanor. 1973. "Natural Categories." *Cognitive Psychology* 4 : 328–350.

Rosch, Eleanor, et al. 1976. "Basic Objects in Natural Categories." *Cognitive Psychology* 8 : 382–439.

Rozin, Paul, and Carol Nemeroff. 1990. "The Laws of Sympathetic Magic." In James Stigler, Richard Schweder, and Gilbert Herdt, eds., *Cultural Psychology: The Chicago Symposia on Human Development,* pp. 205–232. Cambridge: Cambridge University Press.

Ruse, Michael, and Edward O. Wilson. 1985. "The Evolution of Ethics." *New Scientist* 17 : 50–52.

Sanfrey, Alan, James Riling, Jessica Aronson, Leigh Nystrom, and Jonathan Cohen. 2003. "The Neural Basis of Economic Decision-Making in the Ultimatum Game." *Science* 300 : 1755–1758.

(ID CS4 6 September 2010 12:49 PM) UHP (7×10") Optima J-2344 PEW 61:1 (idp) PMU:(idp[KN]/A)6/9/2010 pp. 80–109 2344_61-1_04 (p. 107)

Slingerland, Edward. 2001. "Virtue Ethics, the *Analects,* and the Problem of Commensurability." *Journal of Religious Ethics* 29 (1): 97–125.

———. 2005. "Conceptual Blending, Somatic Marking, and Normativity: A Case Example from Ancient Chinese." *Cognitive Linguistics* 16 : 557–584.

———. 2008. What Science Offers the Humanities: Integrating Body and Culture. New York: Cambridge University Press.

———. (under review). "Early Confucian Virtue Ethics: A Response to the Situationist Critique."

———. (in preparation). Embodied Virtue: Cognitive Science, Virtue Ethics and Early Chinese Thought.

Slingerland, Edward, Eric Blanchard, and Lyn Boyd-Judson. 2007. "Collision with China: Conceptual Metaphor Analysis, Somatic Marking, and the EP–3 Incident." *International Studies Quarterly* 51 : 53–77.

Sober, Elliott, and David Sloan Wilson. 1998. *Unto Others: The Evolution and Psychology of Unselfish Behavior.* Cambridge, MA: Harvard University Press.

Solomon, Robert, ed. 2003. *What Is an Emotion? Classic and Contemporary Readings.* New York: Oxford University Press.

———, ed. 2004. Thinking about Feeling: Contemporary Philosophers on Emotion. New York: Oxford University Press.

Sugiyama, Lawrence, John Tooby, and Leda Cosmides. 2002. "Cross-Cultural Evidence of Cognitive Adaptations for Social Exchange among the Shiwiar of Ecuadorian Amazonia." *Proceedings of the National Academy of Sciences* 99 : 11537–11542.

Sweetser, Eve. 1987. "The Definition of Lie." In Dorothy Holland and Naomi Quinn, eds., *Cultural Models in Language and Thought,* pp. 43–66. Cambridge: Cambridge University Press.

Taylor, Charles. 1989. *Sources of the Self: The Makings of Modern Identity.* Cambridge, MA: Harvard University Press.

Umiltà, Alessandra, Evelyne Kohler, Vittorio Gallese, Leonardo Fogassi, Luciano Fadiga, Christian Keysers, and Giacomo Rizzolatti. 2001. "I Know What You are Doing: A Neurophysiological Study." *Neuron* 31 : 155–165.

Van Norden, Bryan. 2007. *Virtue Ethics and Consequentialism in Early Chinese Philosophy.* New York: Cambridge University Press.

Wallace, Björn, David Cesarini, Paul Lichtenstein, and Magnus Johannesson. 2007. "Heritability of Ultimatum Game Responder Behavior." *Proceedings of the National Academy of Science* 104 : 15631–15634.

Warrington, Elizabeth, and T. Shallice. 1984. "Category-specific Impairments." *Brain* 107 : 829–854.

(ID CS4 6 September 2010 12:49 PM) UHP (7×10") Optima J-2344 PEW 61:1 (idp) PMU:(idp[KN]/A)6/9/2010 pp. 80–109 2344_61-1_04 (p. 108)

Wheatley, Thalia, and Jonathan Haidt. 2005. "Hypnotic Disgust Makes Moral Judgments More Severe." *Psychological Science* 16 : 780–784.

Wilson, Stephen. 1995. "Conformity, Individuality, and the Nature of Virtue: A Classical Confucian Contribution to Contemporary Ethical Reflection." *Journal of Religious Ethics* 23 (2): 263–289.

Wilson, Timothy. 2002. *Strangers to Ourselves: Discovering the Adaptive Unconscious.* Cambridge, MA: Harvard University Press.

Wilson, Timothy, Dolores Kraft, and Dana Dunn. 1989. "The Disruptive Effects of Explaining Attitudes: The Moderating Effect of Knowledge about the Attitude Object." *Journal of Experimental Social Psychology* 25 : 379–400.

Wilson, Timothy, and Richard Nisbett. 1978. "The Accuracy of Verbal Reports about the Effects of Stimuli on Evaluations and Behavior." *Social Psychology* 41 : 118–131.

Wilson, Timothy, and Jonathan Schooler. 1991. "Thinking Too Much: Introspection Can Reduce the Quality of Preferences and Decisions." *Journal of Personality and Social Psychology* 60 (2): 181–192.

Wong, David. 2002. "Reasons and Analogical Reasoning in Mengzi." In Liu and Ivanhoe 2002, pp. 187–220.

Zajonc, Robert. 1980. "Feeling and Thinking: Preferences Need No Inferences." *American Psychologist* 35 : 151–175.

Zimbardo, Philip, Stephen Laberge, and Lisa Butler. 1993. "Psychophysiological Consequences of Unexplained Arousal: A Posthypnotic Suggestion Paradigm." *Journal of Abnormal Psychology* 102 : 466–473.

(ID CS4 6 September 2010 12:49 PM) UHP (7×10") Optima J-2344 PEW 61:1 (idp) PMU:(idp[KN]/A)6/9/2010 pp. 80–109 2344_61-1_04 (p. 109)