Expert moral intuition and its development: a guide to the debate[1]

Michael Lacewing
Heythrop College

Abstract
In this article, I provide a guide to some current thinking in empirical moral psychology on the nature of moral intuitions, focusing on the theories of Haidt and Narvaez. Their debate connects to philosophical discussions of virtue theory and the role of emotions in moral epistemology. After identifying difficulties attending the current debate around the relation between intuitions and reasoning, I focus on the question of the development of intuitions. I discuss how intuitions could be shaped into moral expertise, outlining Haidt's emphasis on innate factors and Narvaez's account in terms of a social-cognitive model of personality. After a brief discussion of moral relativism, I consider the implications of the account of moral expertise for our understanding of the relation between moral intuitions and reason. I argue that a strong connection can be made if we adopt a broad conception of reason and a narrow conception of expertise.

Keywords: intuition; expertise; development; virtue; Haidt; Narvaez

1. Introduction: ringing the changes in moral psychology

The last 15 years have seen an important shift in moral psychology. The discipline was previously dominated by the work of Kohlberg. Just as Piaget argued that cognition develops from implicit understanding to explicit verbalization, Kohlberg advanced a view of moral cognition based on developmental stages measured in terms of conscious moral reasoning. This view of cognitive development has been widely challenged. The continued importance of implicit cognition was recognised early on in moral psychology by Rest (1979), and reaffirmed by the rise of dual process models of cognition generally. A great deal of evidence from across the sub-disciplines of psychology indicates that we make many decisions without conscious deliberative thought (Keil & Wilson 2000; Hammond 2000; Hogarth 2001). In addition to the familiar, controlled, conscious, 'explicit' processes of perceiving, deliberating, and responding, there are automatic, typically non-conscious, 'implicit' processes that influence our thoughts and behaviour in ways of which we are unaware (Bargh & Chartrand 1999; Chaiken & Trope 1999; Dijksterhuis 2010). These implicit processes are often affectively charged, and it is now recognised that emotions – previously neglected, and often considered a source of bias and error – play a central role, not only in our social and moral judgments (Damasio 1994, 1999), but in cognition more generally (Lewis 2009). There are now multiple conceptualisations of the role, nature and importance of moral 'intuitions' (Lapsley & Hill 2008).

The best known of these recent theories is the social intuitionist model (SIM) of Jonathan Haidt (Haidt 2001; Haidt & Bjorklund 2008). We can use, for now, Haidt's definition of a moral intuition as 'the sudden appearance in consciousness of a moral judgment, including an affective valence (good-bad, like-dislike), without any conscious awareness of having gone through steps of searching, weighing evidence, or inferring a conclusion' (2001: 818). Haidt has argued not only for the importance of intuitions, but has also challenged traditional understandings of moral reasoning, and it is on this issue that a number of philosophers and psychologists have taken

[1] Thanks to Jonathan Haidt and Darcia Narvaez for comments, corrections, and the provision of additional references, including forthcoming work. Thanks also to two anonymous reviewers for comments and corrections.

issue with his claims. Much less discussed, by philosophers at least, is the origin and development of moral intuitions, including whether and how they can be shaped.

The aim of this article is to provide a guide to this debate, focusing on two key positions, Haidt's own and that of Darcia Narvaez. Narvaez's work is much less well-known to philosophers, which is our loss, given its force and scope. She has been one of Haidt's most insightful critics, and has led the application to moral psychology of a model of personality that has become much discussed as a new basis for virtue ethics, refreshing and reviving the old idea of virtues as skills (Snow 2010; Annas 2011).

Central to understanding virtues as skills is the claim that virtues involve a sensitivity to moral situations akin to perception (McDowell 1979, 1985; Jacobson 2005; Goldie 2007). Moral responses delivered by this sensitivity are typically emotionally-charged and are non-inferential – so they qualify as intuitions in Haidt's sense. Haidt and this philosophical debate are talking about the same psychological phenomenon, though characterised differently. Here, then, is the connection to this issue's special theme of moral emotions. What follows is an examination of the empirical moral psychology that complements those virtue theories that assign a central role to emotions in moral epistemology. Thus, while I talk of intuitions and expertise, the connections to emotions and virtue should be born in mind throughout.

The structure of discussion is as follows. In §2, I present Haidt's SIM and the debate over the role of moral reasoning in moral judgment. In §3, after noting Haidt's and Narvaez's agreed understanding of virtue in terms of 'moral expertise', I turn to the question of how moral intuitions develop towards expertise, examining Haidt's emphasis on innate factors and Narvaez's account of moral development. In §4, I briefly discuss whether what it is that moral experts 'get right' is completely relative to culture or beholden to something more universal or objective as well. In §5, I discuss the nature and requirements of moral expertise, picking up again the question of the relation between moral intuitions and reason.[2]

My approach is to analyse the arguments and claims of each side to find agreement wherever possible. I will not attempt to reference every claim, as that would prove highly disruptive to the text. So I note now that I have drawn primarily on Haidt (2001, 2010), Haidt & Joseph (2007), Haidt & Bjorklund (2008), Haidt & Kesebir (2010), Graham et al (2013), Lapsley & Narvaez (2004), Narvaez & Lapsley (2005), Narvaez (2008a, 2010), and Narvaez & Bock (in press). All unreferenced claims attributed to Haidt or Narvaez can be found in these sources.

2. Social intuitionism and its critics

2.1 The Social Intuitionist Model of moral judgment

For readers not already familiar with it, here is a brief outline of Haidt's model. Haidt contrasts moral intuitions, as defined above, with moral reasoning, which is understood as a conscious and intentional process of reflective deliberation. This contrast does not entail that moral intuitions are non-cognitive. They are a kind of cognition involving, in particular, the interpretation of

[2] Another topic that has dominated debates in moral psychology is the scope of morality. Haidt argues philosophers and psychologists have been mistaken to think of morality in terms of harm and fairness, the central concepts of traditional consequentialist and deontological theories. In this, he is joined by many virtue theorists, who tend to favour a plurality of 'thick' concepts and welcome the contribution of emotions to moral theory. This package – emotion, intuition, virtue, pluralism – is favoured by the discussion below.

actions, characters, and social situations.[3] The first claim of the model (entitled 'intuitive judgment') is that, in the vast majority of cases, moral judgments are (or result from) moral intuitions. They do not, at least commonly, result from moral reasoning. Instead, moral reasoning is more commonly post-hoc – seeking out reasons and arguments that support the intuition ('post-hoc reasoning'). So people first adopt a view, and then explain and justify it to themselves. People also use moral reasoning to justify their views to others, and to persuade them. But this process typically works not through engaging others' reasoning, but by triggering intuitions in them, which then form their moral judgments ('reasoned persuasion'). Sometimes people's intuitions are affected directly by the intuitions of those around them, without any reasoning being offered ('social persuasion'). The mere fact that someone's social group holds a particular view is itself a significant influence on the views that a person holds.

It is possible to override one's intuitions through reasoning to come up with a contrary moral judgment ('reasoned judgment'), but this is rare, and will usually only occur when someone's intuitions are weak and their ability to reason about the case is very high. (If the intuition is strong, they may end up with a 'dual attitude', i.e. a reasoned judgment that they espouse and an intuitive judgment that unconsciously influences their behaviour.) It is also possible, through reflective deliberation, to trigger spontaneous new intuitions that contradict the initial intuition ('private reflection'), e.g. by adopting another person's perspective, though this is also relatively rare. In these cases, one's judgment may be determined simply by the stronger intuition, or by further reasoning that resolves the conflict.

One of the model's primary attractions is its consilience with findings in social psychology, neuroscience, developmental psychology, and evolutionary psychology. Some of the findings support the view that emotions are centrally implicated in the processes involved in making moral judgments, others that such processes are automatic, others that such processes are unconscious, while still others support Haidt's claims about the limitations of conscious reasoning. A brief account gives a flavour of the research drawn upon.

That moral judgments result from emotional processes is shown by three types of evidence (Prinz 2007, Ch. 1):

1. Neuroscientific: Damasio (1994) shows that people with damage to their ventro-medial prefrontal cortex are unable to integrate emotions into their judgments (see also Koenigs et al 2007). If the damage occurs in adulthood, these subjects, while able to reason normally, are unable to judge (without huge effort) that certain actions should or shouldn't be done. If the damage occurs in childhood, the subjects exhibit similar behavioural patterns to psychopaths, showing moral callousness (Anderson et al 1999). A number of fMRI studies also indicate that moral judgments are correlated with activity in brain regions involved in emotional processes (e.g. Greene et al 2001; Moll et al 2002, 2003).
2. Psychopathy: while psychopaths can apparently use moral concepts, in that they can identify certain actions as right or wrong by the standards of society, they are unable to grasp their moral import, i.e. they fail to distinguish moral from conventional rights and wrongs (Blair

[3] That emotions involve considerable cognition of this kind is widely accepted in psychology (Lazarus 1991), although some question whether the mechanisms involved deem the title 'cognition' (Prinz 2004). But this cognitivism should not be confused with, and does not entail, any form of *metaethical* cognitivism. The facts cognised in the psychological theories are social facts (motives, meanings, etc.); they are not truths about the objective value of features of situations. Metaethical cognitivists and non-cognitivists can agree on this even as they dispute the vexed question of whether moral intuitions and judgments are cognitions of distinct evaluative properties.

1995, 1997). The current leading explanations of psychopathy identify their moral deficit as a result of emotional impairments (e.g. Blair 2007; Kiehl 2006).

3. Social psychology: The manipulation of subjects' emotions changes their moral judgments. Wheatley & Haidt (2005) hypnotised subjects to associate disgust with a neutral word, and found that such subjects made harsher moral judgments of characters in vignettes that contained the associated word. Schnall et al (2008) showed that seating subjects at a dirty desk likewise increased the severity of their judgments, as did asking them to make moral judgments in the presence of a bad smell.[4] Further support comes from Westen's (2007) work on political psychology. In a survey of people's political views, their emotional responses to policies and individuals accounted almost entirely for their judgments while factual knowledge was almost irrelevant.

That the processes yielding moral judgments are automatic and unconscious is again supported by three sources of evidence:

1. Dual process models: the extensive research in dual process models yields this result for other, similar social attitudes and judgements. Thus, many studies show that we make evaluations of people and social situations very rapidly and without being aware of so doing (Bargh & Chartrand 1999; Chaiken & Trope 1999; Dijksterhuis 2010). We may expect the processes behind moral judgment to occur in the same way.
2. Moral dumbfounding: people reach (and hold) moral judgments without being able to give their grounds (Haidt & Hersh 2001). For example, people judge incestual sex wrong, even in a hypothetical case in which it happens once, with contraception, with no harm done to either sibling or their relationship, and in complete secrecy, but they were unable to justify their judgment, citing features such as psychological harm and possible birth defects that are ruled out by the case.
3. Child development: very young children can recognise and evaluate morally good and bad behaviour before they are able to consciously deliberate (Hamlin, Wynn & Bloom 2007; Warneken & Tomasello 2006).

Many of these sources also support the view that these automatic, unconscious processes are emotionally charged. Of course, not all are or need be; the significant contribution of those processes that are affective themselves or produce conscious affective responses is sufficient for Haidt's purposes.

We should, I think, accept that many moral judgments are produced by such emotional, automatic processes. Indeed, the centrality of moral intuitions is no longer much disputed in moral psychology. The debate is whether reason is as ineffective in correcting and directing such intuitive processes as Haidt argues.

2.2. The debate over moral reasoning

I shall not here try to assess the current debate between those who doubt the extent, accuracy, and power of moral reasoning and those who defend it.[5] Instead, I lay out four difficulties attending it. First, both sides tend to be selective in the studies they cite, failing to discuss satisfactorily the evidence for the opposition. Second, significant problems of ecological validity

[4] There is also a large body of literature implicating emotional involvement in moral motivation and action, but as our concern here is moral judgment, we may set that aside. It is notable, however, that Haidt does not always observe this distinction, inappropriately taking such evidence as supportive of his model.

[5] For recent discussions, see Tiberius 2008; Kennett & Fine 2009; Harman, Mason & Sinnott-Armstrong 2010; Huebner 2011; Sauer 2011, 2012.

make it unclear what general conclusions should be drawn from many of the experiments. Third, the debate concerns a matter of degree. Finally, there are significant disagreements about what qualifies as a process or product of reason.

It is important to note that Haidt's theory is, in the first instance, descriptive – a set of claims about what *is* and *can be* the case for the majority of people. It would be a mistake to raise objections starting from a normative model of moral psychology (e.g. how people should reason) if this is only realistically possible for a small minority of people.

2.2.1 Selective evidence

In support of his view of moral reasoning, Haidt presents some of the considerable evidence for *motivated reasoning* (Kunda 1990; Chen & Chaiken 1999; Moskowitz, Skurnik & Galinsky 1999; Ditto, Pizarro & Tannenbaum 2009). People seek to manage the impression they give to others and to ensure interactions with them go smoothly, and they adjust their views and reasoning in light of this. People also alter their reasoning to defend themselves from experiencing cognitive dissonance and from information that threatens their commitments and worldview. More generally, people tend to exhibit *confirmation bias* in reasoning, such that they gather, attend to, and interpret evidence in such a way that it favours the views they hold (Nickerson 1998). Nisbett and Wilson (1977; Wilson 2002) argue that, in making inferences about the causes of one's own and others' behaviour, people's reasoning is *post-hoc*, searching for what *would* make sense of the behaviour, rather than correctly identifying relevant factors. Such rationalisation is also apparent in split-brain patients (Gazzaniga 1985): subjects construct explanations (the verbal centre is in the left hemisphere) for the activities of their left hand (controlled by the right hemisphere) which they are not conscious of guiding. Haidt argues that the same post-hoc processes are at work in moral reasoning – people consult their theories about relevant moral reasons, rather than the actual (unavailable because unconscious) processes that produce the judgments.

Such biases are *not corrected for* by intelligence (Perkins, Farady, & Bushey, 1991) nor by additional training in critical thinking, as this does not appear to transfer reliably from classroom to real-life settings (Nickerson 1994, Willingham 2007). Philosophers are extremely unusual in spontaneously looking for reasons both for and against a position (Kuhn 1991); otherwise, efforts to reduce biased thinking have been disappointing (Lilienfeld, Ammirati, & Landfield, 2009).

But this is all just one side of the story. A number of critics, including Narvaez, have taken evidence of the complex interplay between explicit and automatic processes to indicate reason's input and influence. People are able to *exert a degree of control and correction* over the products and influence of automatic processes through increased motivation to be accurate or unbiased (Monteith et al 2002; Kunda & Spencer 2003), consciousness of accountability (Lerner et al 1998), and careful and reflective thought in general (Wegner & Bargh 1998; Hogarth 2001; Gawronski 2004). Some forms of such goal-directed influence can themselves become automatic (Payne et al 2005; Glaser & Kihlstrom 2005; Gollwitzer, Bayer & McCulloch 2005; Trope & Fishbach 2005). *Education* and the acquisition of new information challenges and changes previous intuitions (Plous 2003). Particular practices and activities can *improve moral reasoning* over time (DeVries & Zan 1994; Power, Higgins & Kohlberg 1989), as can a liberal arts education (Pascarella & Terenzini 1991). Achieving the most developed form of moral reasoning – which Narvaez follows Kohlberg in designating 'postconventional', demonstrating an independence from one's cultural morality – predicts attitudes towards public policy on moral issues better than political or religious attitudes do (Narvaez, et al 1999; Thoma, et al 1999), which would be odd if reasoning had little influence on one's moral views.

What is one to believe? Neither side takes much time to address the evidence marshalled by the other. What follows are three reasons why this might be so.

2.2.2 Ecological validity

Despite powerful defences of his view, Haidt admits outright that 'The precise roles played by intuition and reasoning in moral judgment cannot yet be established based on the existing empirical evidence...' (Haidt & Kesebir 2010: 807). One reason for this is that many of the experiments (including many of those cited above) have involved 'highly contrived' situations, designed to maximise the contrast between intuitive processes and reasoning. Wegner & Bargh (1998) argue that the problem afflicts dual process theories generally, as many of the experiments demonstrating the strong influence of unconscious, automatic processes do so as a result of experimental design. We need to know, in everyday not laboratory situations, how often people revise their intuitive judgments, and what influences such revisions; what factors lead to more deliberative judgments; and when these are better than intuitions.

2.2.3 Matter of degree

Both sides agree that the influence of reasoning, and moral reasoning in particular, is a matter of degree. On the one hand, we should all accept that people often demonstrate motivated, biased, or post-hoc reasoning. On the other, Haidt's SIM allows that 'reasoned judgment' and 'private reflection' occur. But, he argues, this is *rare*. In most situations, we make moral judgments intuitively, and because we have interests and commitments at stake, reasoning, when it occurs, is one-sided. However, he accepts that if private reflection is more common than he claims, his model would need to be altered. Narvaez argues that there is evidence of exactly this (Klinger 1978; Pizarro & Bloom 2003). For instance, we frequently have to assess multiple demands upon us (as colleague, parent, sibling) as well as balance our own needs against those of others. Both sides agree that where intuitions conflict, reasoning has a larger role to play, and that diary studies could provide useful evidence of the frequency of this.

2.2.4 What is reason(ing)?

Haidt argues that most moral reasoning that generates or changes judgments works by triggering new and conflicting intuitions, rather than providing valid reasons and arguments. For example, Kohlberg and Narvaez emphasise the importance of adopting alternative points of view in postconventional reasoning. But this triggers new intuitions, rather than producing anything recognisable as formal reasoning. Narvaez largely agrees, but argues that *seeking out* alternative viewpoints in 'moral imagination' and resolving the conflicting intuitions that arise is not a function of intuition; it is an achievement of reason.

Again, both Haidt and Narvaez emphasise the significant effect of situations on moral judgment. While Haidt notices how our intuitions respond automatically to situations, Narvaez notes that we can deliberately seek out situations that will bring about a change in our intuitions or adopt other means of indirect influence, such as mindfulness training and psychotherapy. But while she counts these in favour of the influence of reason, Haidt (who is equally happy to recommend such courses of action – Haidt 2006) takes the need for such measures to indicate that reasoning does not have much influence.

Are we right to *contrast* intuitions and reasoning? Haidt notes that it is only when we contrast them as *distinct cognitive processes* that we can conduct the sorts of experiments cited above and construct the related models. But if we interpret the idea of reason more generously, as defenders of moral reasoning do, the contrast loses its focus, and models in which reason and

intuition are *complementary and interdependent* suggest themselves: Reasoning is informed by intuitions which, in turn, can be more or less reasonable.

The debate over the scope of reason takes place against the backdrop of another debate concerning evolutionary psychology. Haidt favours a *socio-functional* account of reason. The purpose of thinking is 'doing'; the purpose of moral thinking is to manage our social relations and environment. Because we are 'ultrasocial' creatures (Lerner & Tetlock 2003), these relations contributed crucially to our survival, and reasoning developed to serve this end (Mercier & Sperber 2011). This debate raises larger questions about the nature of reason and the relation between its practical and epistemic uses, and about the reliability and power of evolutionary psychology as a discipline (Panksepp & Panksepp 2000; Dupre 2001). Human identity is at stake here. Evolutionary psychology and dual process models tend to promote a model of the self – the concern of the 'personal' level of psychology – as something 'virtual' (Frankish 2009), a mere 'shell' (Gardner 2000). Many moral psychologists and philosophers are unhappy with this depiction, finding more substance in our abilities to develop and deliberate in rational ways. There is a difficult path to tread between outmoded conservatism that fails to recognise the importance and extent of automatic, unconscious, affective processes and a quintessentially human but potentially dangerous and misleading love of the new and radical.

We have looked at Haidt's basic model and outlined the debate over moral reasoning it has engendered. This debate, which I have not sought to resolve, has focused on the interrelation between intuitions and reason in the psychological functioning of adults. But, given that we must clearly take seriously the place of moral intuitions in moral psychology and moral epistemology, we will want to know how we come to have such intuitions in the first place, and whether they can be shaped during development. What follows focuses on this issue and its implications.

3. Virtue, expertise, and the development of intuitions

3.1 Intuitions and virtues

Haidt and Narvaez agree that the implicit nature of the processes behind moral intuitions does not entail that they cannot be trained or developed: intuitions can be naïve or they can be educated. And they are agreed in framing their accounts of moral development within the virtue tradition. Thus, they agree that intuitions are to be understood in terms of a general moral competence that links together perception, motivation, action, and reasoning in situations which demand a moral response. They also agree that this competence is best understood as a form of expertise. Both emphasise the claim that it does not involve the application of general rules, but a finely attuned sensitivity to the complexities of individual social situations.[6] The development of moral expertise is thus the development of virtue as a cognitive-affective whole, roughly as theorised by neo-Aristotelians such as McDowell and Nussbaum.[7]

In this approach, Haidt and Narvaez stand together in contrast to theorists who, drawing on analogy with other types of automatic unconscious cognitive processes, understand moral intuitions in terms of heuristics (Cosmides & Tooby 2004; Gigerenzer 2008) or a universal moral grammar (Hauser 2006; Hauser, Young & Cushman 2008; Mikhail 2007). In what follows, I

[6] Haidt oversimplifies, then, when he speaks of children having 'crude and inappropriate' evaluative emotional responses until they learn the 'application rules for their culture' (Haidt & Bjorklund 2008: 206). More than learning rules will be needed to 'get it right'.

[7] There are two qualifications here. First, our focus in discussing moral expertise will be in relation to moral judgment, rather than moral action (about which Haidt's model has little to say). Second, we may wish to draw a distinction between moral expertise and virtue in terms of their sensitivity to objective questions of human flourishing – see note 18.

present the case for thinking of moral intuitions and their development within the virtue framework without concerning myself to rebut the alternatives.

3.2 The contribution of nature

The SIM as outlined takes no stance on the origin of intuitions. But Haidt supplements the model with an account that strongly emphasises innate factors. By 'innate', he means only 'organised, to some extent, in advance of experience' (Haidt & Joseph 2007: 367). His account is intended to support the claim that the primary drivers in shaping moral intuitions are non-rational. Innate factors provide an initial structure for moral intuitions, a 'first draft' which is then modified by experience and culture. The processes of modification are themselves governed by innate factors.

Haidt's recent work, with colleagues, has focused on the development of 'Moral Foundations Theory', and the evidential base is developing extremely quickly (Haidt 2012; Graham et al 2013). The theory claims that moral intuitions derive from a number of 'foundations', found across almost all cultures and societies, which comprise our native inheritance. Of these, five command the greatest evidential support, viz. care/harm, fairness/ cheating, loyalty/betrayal, authority/subversion, and sanctity/degradation.[8] There may, however, be others, e.g. liberty/oppression (Haidt 2012) – the list is unlikely to remain fixed as research progresses. For something to qualify as a moral foundation, it should meet five criteria (the names are my own):

1. Moral: it is found in judgments deploying shared norms which people make regarding others' actions, even when those actions do not directly affect themselves.
2. Intuitive: it generates a specific pattern of automatic, affective evaluation.
3. Universal: it is found across most cultures, and most importantly in 'hunter-gatherer' societies, as these most resemble the lifestyles under which human beings evolved.
4. Preparedness: it is found in non-human primates and emerges in young children before they have been taught its importance
5. Advantage: an evolutionary explanation demonstrates its adaptive advantage. Each of the five foundations relates to an adaptive challenge that we may speculate our ancestors faced, such as caring for the young or injured, reaping the benefits of cooperation (within the group and with non-kin), negotiating hierarchical social structures, and avoiding disease. This explanation can be linked to Nussbaum's (1993) idea of grounding the virtues in 'spheres of existence' – areas of life in which all human beings, given their nature, must make choices and develop responses.[9]

---

[8] Prior to 2012, these received different names, viz. harm/care, fairness/reciprocity, in-group/loyalty, authority/respect, and purity/sanctity.

[9] While Haidt's official commitment is only to 'evolutionary preparedness', he favours a modular account of foundations, using Sperber's (2005) flexible concept of a module. On this view, the claim is that moral intuitions are generated by processes that are modular 'to some interesting degree', which may vary from one module to another. The processes underlying the five foundations may be more or less domain-specific and only partially encapsulated, and rather than generate outputs directly themselves, they can be thought of as 'learning modules' – innate learning mechanisms that generate more specific modules over the course of development within a particular culture (disgust may operate like this, such that we learn to become automatically disgusted at certain things). Whether a modular interpretation of the structure of intuitions is right in the end is not something Haidt battles for, even allowing that what is learned are not specific modules but 'bits of subcultural expertise' (Haidt & Joseph 2007: 381), connections between perceptions of certain situations and emotional evaluative responses that cannot be easily controlled or revised.

Graham et al (2013) reviews the evidence for each of the these criteria in relation to each of the five foundations. That they meet the first two criteria is uncontroversial; but there is considerable debate over whether they meet the last three. For example, engaging with the third and fifth criteria, Fry & Souillac (2013), conducting a review of the ethnographic data on nomadic forager societies, find support for the view that care and fairness are foundations, but not loyalty or authority (they do not review sanctity). They argue that concerns with loyalty and authority emerge in the last 10,000 years, after the development of agrarian societies and too recently to receive an evolutionary explanation (see also Narvaez 2013b). Their current widespread occurrence does not, therefore, support their status as foundations with an evolutionary history. On the other hand, Graham et al (2013) cite evidence that precursors of both loyalty (de Waal 1982) and authority (Boehm 1999, 2012) are found in primates, indicating preparedness, and that an evolutionary explanation of their adaptedness can be given. In response, Fry & Souillac note that there is variation among primate species, being more notable in chimpanzees, and less so in bonobos.

Drawing on MacLean's (1990) theory of how the brain evolved, Narvaez offers a different structure for moral psychological theory. She identifies three distinct moral systems – the ethics of security, of engagement, and of imagination, each rooted in distinct brain circuits. The ethics of security is grounded in neurocircuits that deal with fear, anger, and sexuality (the 'R-complex' (MacLean 1990) or 'extrapyramidal action nervous system' (Panksepp 1998)). It typically operates implicitly, and is concerned with self-preservation and personal gain, including status and loyalty. The ethics of engagement is based in neurocircuits that deal with social engagement and cultural membership (Panksepp 1998; Nelson & Panksepp 1998; Schore 2003; Greenspan & Shanker 2004). It is concerned with meaningful relationships with others. The ethics of imagination is grounded in the neocortex, especially the prefrontal cortex. Its primary work is that of deliberation, of coordinating the concerns and intuitions of the other ethical systems. Through learning and willpower, it is able to shape which stimuli generate emotional responses; and it develops a personal narrative identity within a social narrative adopted from the social environment.[10]

Narvaez is sceptical about modules. While we have many specialized subcortical networks, many of which we share with other mammals, 'there is no comparable evidence in support of highly resolved genetically dictated adaptations that produce socio-emotional cognitive strategies within the circuitry of the human neocortex' (Panksepp & Panksepp 2000: 111). The idea of sophisticated cognitive modules has no basis in neuroscientific fact. The neocortex has a high degree of plasticity, even as it is grounded and structured by the propensities of subcortical processes (Panksepp 1998). Narvaez's interpretation is to think of the innate pre-structuring of moral intuitions in terms of units resulting from *experience* structured by this combination of subcortical adaptations and neocortical plasticity.

Given how loosely Haidt is willing to use the term 'module' in this context, perhaps this alternative is really a matter of emphasis. As Panksepp (1998) shows, emotions and concerns are substantially underpinned by subcortical adaptations. Importantly, on either model, there is some innate structuring.

[10] Many intuitions related to loyalty and authority relate to the ethics of security, while many intuitions related to care and fairness map onto the ethics of engagement. The ethics of imagination defends the role of moral reasoning about which Haidt is sceptical. It is interesting to note that it involves what we may characterise as balancing four of Haidt's five foundations – Narvaez talks of reciprocity between the law and the individual (balancing fairness and care with authority) and coordinating community and individual interests (balancing fairness and care with loyalty).

Once again, rather than adjudicate the debate, I want to emphasise the similarity between these viewpoints. The conclusion to draw is that whether intuitions are prestructured by five foundations or two ethics (or even five-into-two), they are prestructured. This is what we would expect on any virtue theory with an Aristotelian flavour. Human nature provides both the limitations and the content of morality (Foot 2001, Hursthouse 1999). What implications does this have for how intuitions can be shaped and developed? We turn to the question of expertise.

3.3 The development of expertise

For Haidt, mature moral functioning is not a matter of *simple* intuitions, but ones structured by virtues. This involves a form of expertise in perceiving morally relevant information and responding appropriately that is, in part at least, learned or acquired, involving significant extensions in the number, variety, conceptual complexity, and sophistication of one's intuitions. However, Haidt understands moral development as 'primarily a matter of the maturation and cultural shaping of endogenous intuitions' (2001: 828). Children are set up to have intuitions grounded in the five (or more) foundations, and moral development involves 'assisted externalization' (Fiske 1991). Like language and sexuality, morality emerges on a developmental timetable through interaction with a local cultural environment. At various points of development, children will show concerns and emotions related to a specific foundation. They quickly come to recognise and conceptualise a wide diversity of social situations in these terms, and react with automatic emotional evaluations. Socialisation is a matter of fine-tuning, fixing the extension of concepts, developing specific forms of sensitivity. Culturally-specific moralities are formed by elaborations and specifications of the foundations towards particular virtues and through selective loss (not all cultures emphasise virtues related to loyalty, authority, and sanctity).

Narvaez objects that Haidt's account is unsatisfactory, and he has come to accept her criticism that Moral Foundations Theory has focused insufficiently on the question of how moral intuitions develop. In particular, no account is offered of development of expertise and individual differences in it.[11]

Built into Narvaez's model is a hierarchy of moral concerns and emotions, which she draws upon in her account of moral development. It is a hierarchy that Haidt repeatedly emphasises reflects the distinctive morality of educated liberals, and not one he believes can be grounded in fact. Of course, we needn't accept the hierarchy along with the neuroscience, but Narvaez (2013c) provides evidence that connects the security ethic to poorer psychological functioning, in particular, insecure attachment and the use of aggression or withdrawal in social interaction, while Wright & Baril (2013) argue that it is connected to psychological defence (discussed below).

[11] Haidt discusses individual differences in moral judgment, emphasising not what is learned, but what is innate. Twin studies have shown monozygotic twins are more similar than dyzygotic twins, and that monozygotic twins reared apart are almost as similar as those reared together (Bouchard 2004). Given this, we can expect innate temperament to play an important role in moral psychology. So Haidt speculates that the strength of intuitions deriving from the five foundations will differ between the foundations within one person and will differ between people for the same foundation. Again, given that children are differentially responsive to reward and punishment (Kochanska 1997), they will respond differently to attempts to 'tune up' their intuitions through socialisation. Some may be more prone to social persuasion than others. Differences in cognitive ability will be reflected in differential use of the four forms of moral reasoning (post-hoc, reasoned persuasion, reasoned judgment, private reflection). All this gives us difference, but not in expertise.

Narvaez situates the development of moral expertise in the context of the development of the self, and adopts and develops a popular social-cognitive model of personality to make the case.[12] Mischel & Shoda's (1995) CAPS (Cognitive-Affective Processing System) model posits that personality traits are 'social-cognitive units', consisting of integrated structures of beliefs, desires, feelings, goals, expectations, values, and plans. These units are sensitive to situational features and to other psychological states and events. Their integration means that the activation of one element leads to the activation of others. For instance, compassion involves being moved to sadness (feeling) by the misfortune of others (belief/experience), being motivated to help (desire), and evaluating such help positively (value). Thus, the model posits the activation of mental representations (of social situations, oneself, others, prospective events, goals and so on) in personality functioning, such that social-cognitive units influence perception, interpretation, and response in social situations – precisely what is at stake in moral expertise.

Central to such representations are schemas, the 'cognitive carriers of dispositions' (Cantor 1990: 737). Schemas are 'organized knowledge structures that channel and filter social perceptions and memories…. and guide our appraisal of social situations, our memory for events, and our affective reactions' (Lapsley & Narvaez 2004: 195). (Lest the connection with emotion gets lost, it is worth repeating that many of the relevant cognitive processes and products involve emotion.) Schemas that underpin our personality traits are chronically 'accessible', i.e. the information structured in the schema is available for (conscious or unconscious) cognitive processing. Schemas explain the well-established connection between personality and the interpretation of social situations,[13] and much of the evidence supporting dual process models indicates the automatic, unconscious activation and influence of processes involved in social-cognitive units.

Schemas and their accessibility are generally thought to emerge from frequent and consistent *experience*, from the developmental history of the individual (Bargh, Lombardi & Higgins 1988). Narvaez utilises the idea of a 'general event representation' (Thompson 1998). Repeated events, such as family routines and rituals, give rise early on to abstracted or generalised representations of social interactions, which develop further into 'working models' of social experience, including one's role and what to expect. These can be thought of as a child's first schemas. Further interaction and dialogue with caregivers contributes to the construction of autobiographical memories and personal narratives, integrating schemas into the child's personality. This model of personality development has been independently arrived at by

---

[12] The model derives from the work of Mischel and Shoda (Mischel 1968, Mischel & Shoda 1995, Cervone & Shoda 1999), and has become widely accepted in the wake of the situationism-character debate. That debate considered evidence that people's behaviour is so malleable by situational influences that the existence of traits of character is called into question. Social-cognitive models have been a popular solution, holding that behaviour is the result of a dynamic interaction between genuine features of the person and the situation. Once situations are categorised by how the subject interprets them, then there is strong evidence of stable and predictable individual behavioural responses across diverse situations with similar meaning. For an account of the debate and detailed defence and elaboration of the model, see Snow (2010).

[13] For example, aggressive people are more likely to automatically notice and recall hostile cues (Zelli, Huesmann, & Cervone 1995), authoritarian people are more likely to automatically infer others are authoritarian from prompts that are ambiguous (e.g. enjoying military parades) (Uleman et. al. 1986), and how one understands a moral narrative is strongly influenced by one's chronically available moral schemas (Narvaez 1998; see also Narvaez et al 2006).

theorists drawing on distinct bodies of evidence (e.g. Stern 1985; Cassidy & Shaver 1999; Andersen & Chen 2002; Andersen & Thorpe 2009).[14]

As the child becomes older, schemas develop in response to wider social settings with peers, teachers, and so on. The process, argues Narvaez, models the development of expertise (Dreyfus & Dreyfus 1990). It is initially difficult for the child to perceive what is required and how to respond, but with time, immersion in examples and opportunities to practice, and constant (implicit and explicit) feedback, schemas are enriched and become chronically accessible. The child begins to attend to relevant facts and develop the necessary skills to negotiate social challenges successfully. Rule-bound procedures practised in specified settings yield to an integration and flexibility in response, which becomes increasingly automatic. Thus early experience establishes the structure and shape of one's intuitions and reasoning in various domains, including the social (Schore 2001), cognitive (Greenspan & Shanker 2004), and moral (Narvaez 2008b).

Expertise will be a matter of the degree to which schemas (and the social-cognitive units they underpin) have developed and become chronically accessible. Narvaez picks out three ways in particular that experts differ from novices, each of which can be explained in these terms. First, experts have richer, more organized knowledge of the relevant domain than the novice (Abernathy and Hamm 1995; Sternberg 1998, 1999). Such knowledge is embedded in more richly developed schemas. Second, experts perceive the world differently, picking up relevant details and opportunities that novices miss (Johnson and Mervis 1997). Because schemas influence social perception and interpretation, experts with well-developed, chronically accessible schemas notice what novices don't, and do so automatically. Third, experts solve problems effortlessly and automatically. Experts know better what to do and do it more automatically than novices as chronically accessible schemas make the relevant information immediately available and because they have developed the routinized repeated behaviour that accompanies such schemas (Ericcson and Smith 1991; Feltovich, Prietula, & Ericsson, 2006).

The intuitions of *moral* experts, then, have been developed and shaped through the development of schemas for understanding the social world, so as to have achieved a richer moral knowledge, a greater sensitivity to morally important features of situations, and greater fluency and appropriateness of response than those who are not moral experts. These developments have taken placed as part of the formation of their personality (or character, as virtue ethicists have it).

Given that Narvaez's account is compatible with the influence of innate factors, I cannot see anything in the account so far that need be objectionable to Haidt.[15] Of course, there remains a

---

[14] Narvaez (2013c) relates moral development to her three 'ethics'. I have left out this additional detail to avoid over-complication. In brief, she argues that whether an individual acquires a dominant security ethic or a dominant engagement and/or imagination ethic is heavily influenced by interpersonal relations are in early childhood. The former is the 'default' when nurturance is poor, while the latter develops when children receive warm, responsive care. On her account, experts will, of course, need to develop an imaginative ethic. This line of evidence relates to remarks made at the end of §4 and the discussion of defence (which is correlated with insecure attachment) in §5.4.

[15] Haidt's views on the relation between moral development and the self have shifted. Haidt & Joseph (2007) hold that what is learned are the skills of perception and response. Haidt, Graham & Joseph (2009) and Graham et al (2013) adopt McAdams (1995) three-level model of the self, but, as noted, accept that their account requires development. Level I is very abstract traits (the 'Big Five'). Level II involves 'personal concerns' – the kind of thing McAdams has in mind

difference of emphasis on 'nature' and 'nurture'. Nevertheless, Haidt needs a theory of how moral expertise develops, sensitive to local cultural demands, and this one draws on many very well-grounded and independent claims.

With this account of the development of expertise in place, we may next ask what the implications are for our understanding of moral intuitions. Experts must, in some meaningful sense, 'get it right' – or there is no such thing as expertise. What is the status of what they 'get right'? What criteria for reaching this judgment are available? What is the relation between moral expertise, reason, and anything that could merit the moniker 'moral truth'? In the next section, we briefly consider this issue from the perspective of comparing or improving upon cultures' moralities. In §5, we discuss criteria for rational functioning.

4. Relativism or pluralism?

Even on the most deterministic reading of Haidt's developmental story, there is sufficient flexibility in the way people develop for us to make sense of the question 'are there better and worse ways to develop, better and worse sets of intuitions?' If there are, then it is normatively important to try to identify the better ways, the set (or sets) of intuitions cultures should seek to inculcate.

Haidt follows Wiggins (1987) in holding that moral truths are 'anthropocentric truths' – they hold true only for the sorts of creatures human beings are. Haidt also wants to retain a naturalistic ethics, in which morality is grounded in facts, in particular, facts about human psychology. These facts, Haidt argues, indicate that we must adopt pluralism – not only is there more than one source (foundation) of moral value, values are incommensurable, such that there is no single best combination or hierarchy of them.[16] All three claims are widely held, at least among virtue theorists, and I see no reason for Narvaez to dispute this metaethical picture so far.

Haidt then rejects relativism as 'going too far',[17] but he gives two distinct reasons for doing so. In Haidt & Joseph (2007: 387), he argues that virtues are not completely culturally relative because, as Nussbaum (1993) elaborates, they involve 'functioning well in a specific "sphere of existence"'. 'Functioning well', is understood by Nussbaum, following Aristotle, to be robustly *objective*, even if open to multiple cultural interpretations and shot through with evaluative thought. Thus an entire culture could be mistaken about whether their morality enabled them to function well. In Haidt & Bjorklund (2008: 215-16), he rejects relativism because 'moral systems can be judged on the degree to which they violate important moral truths held by members of that society… A well-formed moral system is one that is endorsed by the great majority of its members', including those who are 'victims' or worse off from an outsider's point of view. This second criterion is *subjective*, in that if the great majority of people within a society believe their

---

(1995: 376) maps closely on to the specific contents of schemas and social-cognitive units, demonstrating that Narvaez's account can, to a significant extent, be made compatible with that of McAdams'. Level III concerns autobiographical narrative. However, McAdams' levels are defined *epistemologically* (the terms in which we may know a person), while Narvaez provides an account of how personality *functions*. For this reason, I would argue that her account is more satisfactory.

[16] Part of the reason for Haidt's reticence on the development of expertise is his general reluctance to talk of 'better' and 'worse' in the domain of moral psychology.

[17] Narvaez is thus mistaken when she accuses Haidt of adopting moral relativism. At times, he does appear to define virtue in these terms, e.g. 'a fully enculturated person is a virtuous person' (Haidt & Bjorklund 2008: 29) – as though it were sufficient for virtue to adopt whatever norms one's culture holds. But his official position is one of pluralism, not simple relativism.

shared morality is correct, then there is nothing further to be said. Put another way, it replaces 'functioning well' by a shared belief in functioning well. Such a belief may be an excellent epistemic marker for good functioning, but it is defeasible and does not amount to the same thing.

Given Haidt's extensive appeal to human nature (via evolutionary psychology), it is not surprising that he has most recently given an explicit endorsement of the first reading. Graham et al (2013) ends by approvingly citing Isaiah Berlin's (2001) point that 'pluralism is not relativism—the multiple values are objective, part of the essence of humanity rather than arbitrary creations of men's subjective fancies'. This objectivity is also strongly indicated by Haidt's (2006) work relating contemporary moral psychology to teachings on wisdom from around the world. It is no secret that many of these teachings are counter-intuitive. What the great majority of people *think* will make life go well will in fact make it go worse than it could and vice-versa (e.g. teachings on generosity, forgiveness, money v. time, etc.). In this literature, virtue is certainly not what people tend to *do*, and possibly not what they *think* they should do (at best, people have a 'dual attitude' toward wisdom teachings). Virtue is what will, *in fact*, help them flourish. Should we say that the morality of wisdom is shared by the great majority, since many people in many cultures recognise it as wisdom, not foolishness? Or should we say that it stands as a critique of the shortcomings of majority morality, since the virtues of wisdom are rare? Either reading challenges any simple identification of virtue with what (let alone how) the majority think, and Haidt argues on empirical grounds that wisdom offers a better moral system.

Narvaez (2013a) adds the thought that wisdom can be understood developmentally as mature moral functioning. If she is right, this provides a connection between human flourishing and psychological functioning. If this can be made out in objective terms (e.g. that derive from developmental pathways), the criteria for advocating some intuitions and structures of intuitions over others are strengthened (though given human psychological and cultural complexity, this is unlikely to rule out pluralism).

In sum, cultures can elaborate and structure intuitions in ways that are better or worse, depending on whether the morality of the culture enables or hinders human flourishing and mature psychological functioning. A great deal of what moral experts are sensitive to is relative to the culture within which they live, and standards for appropriate response will likewise heavily depend on specific cultural values. But virtue also involves having intuitions – sensitivities and reactions – that contribute to human flourishing.[18] If Haidt and Narvaez are right about the universality of certain moral concerns and moral wisdom, then environments in which virtuous intuitions can be developed are very widespread. But children may grow up in environments or cultures that make it very difficult to form virtuous intuitions, and cultures can have moral 'blind spots' – certain areas of social life (or spheres of existence), an obvious historical example being the status of women, in which common thought and practice hinders flourishing. But given pluralism, this judgment cannot be reached easily.

[18] At this point, we may wish to make a distinction between moral expertise that is expertise *only* in terms of the local culture's values and virtue, which is additionally sensitive to truths about human flourishing. Narvaez and Haidt don't draw this distinction, and equate the two as we have done thus far. Even allowing for the distinction, epistemologically it is the more helpful to take the case of moral expertise as virtue as the basic case, and explain why moral expertise does *not* amount to virtue in the instances in which it falls short.

5. Expert intuitions, reason and imagination

We need to end by considering the implications of the model of the development of expertise for our understanding of moral intuitions and, in particular, their relation to reason. But before we can do that, we need to specify the scope of the term 'expert' in this context.

We might say that just about every native speaker of a language is an expert in that language. Or that experts are only those with a much deeper understanding of how language works, a greater vocabulary and sense of nuance. Likewise, is anyone who can *grasp* many of the morally significant concepts and distinctions of their culture a moral expert, or should we restrict the term to reasonably *good* people, or should we restrict it further to the wise and moral *paradigms*? There is no right answer, but we need to fix the term. Both Haidt and Narvaez identify moral expertise as the telos of moral development, but Haidt specifies no more than what we may expect of every normal functioning adult, while Narvaez describes the achievement of a small minority (though potentially widely available). As noted above, we do not want a moral psychology that describes the functioning of only a minority of people if this is unrealistic for the rest. That said, theories of psychological functioning (from vision to memory to affective self-regulation) assume an idea of *good* functioning, and we can legitimately reference goals of development in constructing theories. Because moral psychology has normative significance for who we are as people and how we live, developmental goals acquire additional force. In what follows, I assume Narvaez's more narrow sense of 'expertise', noting where a more inclusive interpretation yields a different picture.

What implications does our model of expertise have for our understanding of the relation between intuitions and reason? I discuss the extent to which we may find four features of reason in the development and achievement of moral expertise. My aim, as usual, is to find what common ground there may be between Haidt and Narvaez.

5.1 Consciousness

In defending his theory of 'assisted externalization', Haidt argues that (most) moral intuitions were not first learned consciously to later become automatic.[19] But Narvaez's theory of the development of moral expertise can concur: much of what people learn in developing schemas (and other forms of implicit knowledge) is picked up unconsciously (Ciancolo, Matthew, Sternberg, & Wagner, 2006). The vast majority of non-school learning involves the 'nonintentional, automatic acquisition of knowledge about structural relations between objects or events' (Frensch 1998: 76). And so it may be with learning about morality. That said, a good deal of the learning that shapes moral intuitions clearly does involve conscious correction and reason-giving. I may never have been explicitly told not to have sex with my siblings (Haidt's example), but I was certainly told a great deal about honesty and lying, loyalty and friendship, and any number of other morally important things.

Because schemas – and thus intuitions – develop through both conscious and unconscious processes, what someone may know exceeds what they can verbalise. Peter Railton (forthcoming) has argued that Haidt's case of moral dumbfounding in the example of harmless incestual sex (§2.1) can be explained this way. People's intuitions are reacting, accurately, to the *risk* of psychological harm. Because no harm occurs, Haidt assumes that we cannot object to the act on the grounds of harm. But (people's intuitions tell them) it can be immoral to risk harm even when, as it happened, no harm occurred. Articulating this intuition, however, is difficult for

[19] He also suggests that if one were to claim this, that would entail that *any* intuitions could be teachable. But this doesn't follow: what can be acquired from experience may nevertheless be shaped and constrained by what is innate, so that not any intuition can be developed.

most people to do; but this inability doesn't necessarily cast doubt on the rational basis of the intuition (Sauer 2012). On this view, we should reject Haidt's restriction of reason to conscious reasoning.

5.2 Self-direction

If we use 'expertise' in the inclusive sense, then there are some forms of expertise, such as speaking one's native tongue, whose development does not require much directed effort by the learner. But for some types of expertise, and in almost all cases of expertise in the exclusive sense, self-direction and commitment assume a greater role. From cookery to chess to meditation to philosophy, developing expertise generally requires great dedication on the part of the individual, and other goals must be set aside. Eventually, however, there is little inner struggle about taking up the activity each time. Narvaez identifies 'ethical focus' as an important part of moral expertise – moral goals must be prioritised, and later become chronically accessible. Such commitment is important for the development of one's sensitivity to morally relevant features of social situations, the appropriateness of one's responses, and the clarity of one's moral reasoning. If this is so, then the development of moral expertise, and perhaps moral development in general, is proactive and autopoetic (Juarrero 1999; Varela 1999). Guiding and holding oneself to a path of action for the sake of an end that one values is a form of autonomy. Thus reason as self-direction plays a role. Furthermore, as these developments in sensitivity and judgment involve refining or extending the content one's intuitions, then reason plays this role in the forming the structure of expert intuition as well.

5.3 Moral reasoning again

We noted in §2.2.4 that the most developed, 'postconventional', form of moral reasoning involves seeking out and successfully imagining alternative perspectives (Rest et al 1999). Both Haidt and Narvaez accept that the process involves triggering new intuitions. But then intuitions from different perspectives will be in conflict, and subjects must deliberate to resolve the conflict. The better one is at doing this, the better one's evaluations of the current situation become. So moral experts generate, in imagination, conflicts of intuition which feed into processes of deliberation. Or better, processes of deliberation involve such acts of imagining. Objections regarding the absence of recognisably formal reasoning fall away, since moral reasoning is not primarily a matter of formal reasoning (a view shared by friends of Aristotelian *phronesis*).

This last point provides the basis for responding to Haidt's objection (to theories that emphasise moral reasoning) that critical thinking is hard to teach and transfer across context. As Narvaez notes, the type of imagination involved in postconventional thinking develops through *life experience* and it is not a *general* ability (to imagine alternative viewpoints in any context), but domain-specific (Feshbach 1989, Selman 2003).[20] In light of this, both Haidt and Narvaez can agree that educationally, what is needed for developing moral expertise are certain sorts of guided experiences (Narvaez 2005, 2013c). This is what we would expect, given the account of moral expertise as integrated into one's personality via social-cognitive units, as personality is not taught but developed (from innate predispositions) through experience.

A second objection Haidt makes is that our confirmation bias is best corrected by dialogue with others. But as with imagining alternative possibilities, this involves, but does not reduce to, triggering conflicting intuitions. There are issues around whether we seek out others' conflicting

[20] Compare: a similar ability to imagine future positions in chess is necessary for chess expertise, its development requires a great deal of practice not mere teaching, and it does not transfer to a general ability to imagine future scenarios in other domains.

intuitions, how we respond to them, and how we incorporate them into a coherent overall evaluation. One needs to be wise enough to seek dialogue with others and learn from it. I can see no theoretical reason for Narvaez to reject the importance of having expertise *distributed between* people, rather than located within one person, on many occasions.

Finally, Haidt objects that the psychology is unrealistic – people don't generally seek alternative points of view. But the fact that only a minority of people achieve postconventional thinking does not constitute an objection on its own. However, it is important that the psychology involved be grounded in more widespread developmental processes and not be beyond the possibilities of normal development.[21] It must be granted that the model of expertise developed by Narvaez (and augmented below) has a strong element of the 'ideal' to it. However, Narvaez's specification of expertise involving the imaginative abilities she defends rests on evidence that this achievement can appropriately be considered a (normal) psychological development in moral thinking.

It is *developmental* as it involves the subject's being able to resolve conflicts between intuitions in a more satisfactory way and having greater sensitivity to features of situations that are morally important. It remains *realistic* by postulating that human moral psychology has a normative dimension built into it that invites and defines development beyond the norm. In his (2006, 2012) commentaries on the mutual misunderstanding of different moral systems, Haidt appeals for people to adopt an openness to alternative viewpoints, to attempt to imagine things from the others' perspective. If he is not being falsely optimistic in these popularising books – if he thinks his appeal is not beyond people's abilities to respond – then he too appeals to a form of wisdom that moves beyond the conventional (culturally specified) to the postconventional. I suspect that moral expertise is not something that many can fully achieve, but it remains something towards which it is meaningful for many to strive. It characterises a continuum of development, along which everyone can progress.

Haidt may object that the defence offered here moves away from his socio-functional account of moral reasoning towards an epistemic one, in which reason aims at truth, rather than the management of our social relations and environment. This takes us to our final feature of reason.

5.4 Accuracy (aka 'truth' and 'appropriateness')

Intuitions involve cognition of social meanings and facts in specific situations (a reminder: metaethical cognitivism is not entailed). Thus, one may hold that harm is wrong, but fail to have the intuition that *this* action is wrong because harmful, because one doesn't realise the harm it will cause (ignorance of consequences). One may fail to condemn an instance of cruelty, because one fails to understand the agent's motive (ignorance of motives). One may condemn a kindly word of advice as belittling because one fails to understand it as supportive (ignorance of meanings). And so on. Moral expertise, we have said, involves the ability to read a situation quickly and accurately. An expert's intuitions are attuned to, rather than ignorant of, what is morally relevant here and now. Experts are also sensitive to what they do not know, and willing and able to suspend judgment until further information is available (Hogarth 2001). We have assumed, not unreasonably, that intuitions that are more accurate in these ways are better than those that are inaccurate. So accuracy is a goal of development and, in this way, the intuitions of moral experts are shaped by reason.

[21] Given the difficulties of developing moral expertise, there is a further question of how best to encourage moral behaviour (education? public dialogue? manipulating situations?). But that is a *further* question, and should not be confused for an account of the nature of moral intuitions and expertise.

With failures of knowledge come failures of response – both of emotion and of action (Trout 2009). Thus accuracy feeds into appropriateness of response. Intuitions are themselves emotional responses, and the appropriateness of such responses is, in part, a matter of accurate understanding of the situation to which they are responses, if we extend this to understanding the cultural availability of certain responses and how they will be received.

However, Haidt rejects the idea that moral intuitions aim at the truth, holding instead that they aim at enabling the agent to succeed in the social order. Furthermore, the two goals conflict, so there are many cases 'where distorted perceptions may be more useful than accurate ones' (Haidt & Kesebir 2010: 814). Haidt intends his appeal to social success to explain the *evolutionary origin* of moral thinking. Thus we have an explanation of why people frequently fail to be accurate. But to object to accuracy of the kind just outlined as a legitimate goal for intuitions, for social success to provide *the* criterion for moral expertise, Haidt must adopt a disturbing picture of the virtuous agent as one who manages their relations with others without regard to biased, hypocritical, or inaccurate thinking.

There may be confusions around the scope of 'expert' again. In the inclusive sense, most people are moral experts, i.e. virtuous (enough), managing their social relations well (enough), and yet most people are prone to inaccuracies in their representation of social situations – and these inaccuracies may even contribute to how they are able to manage their social relations. But in the exclusive sense, we expect experts to excel at sensitivity to the social facts *and* at managing social relations.

Fortunately, holding accuracy as a criterion of expertise does not commit one to the objectionable claim that such accuracy should be pursued at the cost of succeeding in one's relations with others; the moral expert knows the value of each and in each case. But what is entailed is that, on the whole and at least for experts, an accurate appreciation of social situations will at least not hinder, and will generally contribute positively to, success in one's relations with others. Thus, for the view to be plausible, we must be able to interpret the evidence that distorted perceptions are more useful in line with this theoretical commitment. There is reason to think this can be done.

For example, Taylor and Brown (1988, 1994) famously argue that people who harbour positive illusions about themselves and others do better than those who do not. But, even if their conclusion can be established, which is uncertain (Badhwar 2008), it is very important that the illusion is mild, as people whose estimations either of themselves or others are more seriously inaccurate do less well. It is also not established whether those holding positive illusions do so from ignorance, a cultivated mild self-deception, or choice. Only the former would even be in tension with, let alone undermine, holding accuracy as a criterion for moral expertise.

Further evidence that accuracy (within limits) contributes to, rather than detracts from, social success comes from studies of psychological defence.[22] 'Defence' refers to psychological means taken to avoid psychological pain, such as anxiety, fear, guilt, shame and envy, particularly in relation to questions of self-esteem (Vaillant 2000; Cramer 2006). There are many different

[22] Chen and Chaiken (1999) identify three motives that drive cognitive processing: accuracy, defence, and impression management, and argue that accuracy is frequently overridden by the other two. Haidt discusses 'impression management' at length, but says little about defence and its impact on our moral intuitions. In raising the matter of defence here, we will have covered all three motives.

forms of defence, and they typically work through unconscious, unintended distortions in understandings and experience of the self, of passions and thoughts, of others, or of the world. There is a good correlation between the use of more distorting (less 'mature') defences and poor psychosocial adjustment (Vaillant 1993: 132, Table 4).[23] And there is a good correlation between the use of defence and psychological disorders (Cramer 2006: 235-6, 253-4), and of course, between psychological disorders and poor psychosocial adjustment.

Moral intuitions can embody defensive thinking. In such cases, the intuition is a misapprehension of the situation to which it is a response. But it also obscures the painful psychological states that the situation arouses, leading to a misunderstanding of oneself as well. This is an important issue for the development of moral expertise – though it has so far failed to inform the debate – as defence mechanisms occur far more commonly than we might think. They are used universally in childhood and adolescence, as they are entirely necessary in psychological development. In forming a healthy self-esteem, there are many small battles in which children must first impose their wishes on their experience of reality and later relinquish such influence, maturing in themselves and their understanding as they go. The task of development involves giving up such distortions in early adulthood, but few people do so completely.[24] (Defences remain adaptive in certain situations in adulthood, e.g. in situations of unbearable conflict or sudden change – a clear example is their use in the grieving process.) The use of defence mechanisms is particular prevalent in conditions of stress and especially when self-esteem is at issue – both of which commonly apply in morally challenging situations. The development of moral expertise, then, will also involve achieving greater maturity in one's psychological defences, moving from those that distort understanding to forms of coping that do not (Lacewing forthcoming).

Reason as aiming at accuracy – even if the accuracy is embedded not in conscious reflection but in unconscious learning and response – plays a role in the development of expert intuitions. Holding this doesn't commit us to the doubtful claim that moral thinking evolved *in order* to represent truths. We can, therefore, agree with Haidt that 'thinking is for doing', having secured that thinking accurately generally contributes to doing well.

Conclusion

In this article, I have provided a guide to some current thinking in empirical moral psychology on the nature of moral intuitions, focusing on the theories of Haidt and Narvaez. After identifying difficulties affecting the current debate around the relation between intuitions and reasoning, I focused on the question of the development of intuitions. Following Haidt and Narvaez, I worked within the framework provided by virtue theory, looking at how intuitions could be shaped into moral expertise and outlining Narvaez's account in terms of a social-cognitive model of personality. After a brief discussion of moral relativism, I considered the implications of the account of moral expertise for our understanding of the relation between moral intuitions and reason. In sum, I argued that a strong connection can be made if we adopt a broad conception of reason and a narrow conception of expertise.

---

[23] This is, however, complicated by IQ - see Cramer 2006: Ch. 8.

[24] Vaillant (1993: 132, Table 4) notes that of those in the top 20% on a scale of psychosocial adjustment at 65 years old, 50% still use less than mature defences, and the percentage for those lower on the scale is considerably higher. Cramer (2006: 204) notes that neurotic defences are likely to survive into adulthood, and remarks on the widespread distribution in 'normal' samples of characteristics defining psychological disorders, e.g. depressive tendencies, phobias, pathological aggression, antisocial traits, etc. that are associated with the use of defence (224, 235).

References

Andersen, S. M. & Chen, S. (2002). The relational self: An interpersonal social-cognitive theory. *Psychological Review*, 109, 619-645.

Andersen, S. M. & Thorpe, J. S. (2009). An IF-THEN theory of personality: Significant others and the relational self. *Journal of Research in Personality*, 43(2), 163-170.

Anderson, S. W., Bechara, A., Damasio, H., Tranel, D., & Damasio, A. R. (1999). Impairment of social and moral behavior related to early damage in human prefrontal cortex. *Nature Neuroscience*, 2, 1032–1037.

Annas, J. (2011). *Intelligent Virtue*. Oxford: Oxford University Press.

Badhwar, N. (2008). Is realism really bad for you? A realistic response. *The Journal of Philosophy*, 105, 85–107.

Bargh, J. A., & Chartrand, T. L. (1999). The unbearable automaticity of being. *American Psychologist*, 54, 462–479.

Berlin, I. (2001). My intellectual path. In H. Hardy (Ed.), *The power of ideas* (pp. 1–23). Princeton, NJ: Princeton University Press.

Blair, R.J.R. (1995). A cognitive developmental approach to morality: investigating the psychopath. *Cognition*, 57, 1-29.

Blair, R.J.R. (1997). Moral reasoning and the child with psychopathic tendencies. *Personality and Individual Differences*, 26, 731-9.

Blair, R.J.R. (2007). The amygdala and ventromedial prefrontal cortex in morality and psychopathy. *Trends in Cognitive Sciences*, 11, 387–392.

Boehm, C. (1999). *Hierarchy in the forest: The evolution of egalitarian behavior*. Cambridge, MA: Harvard University Press.

Boehm, C. (2012). *Moral origins: The evolution of virtue, altruism, and shame*. New York: Basic.

Bouchard, T. J., Jr. (2004). Genetic influence on human psychological traits: A survey. *Current Directions in Psychological Science*, 13 (4), 148–151.

Cantor, N. (1990). From thought to behaviour: "Having" and "doing" in the study of personality and cognition. *American Psychologist*, 45, 735-750.

Cassidy, J. & Shaver, P. R. (Eds.). (1999). *Handbook of attachment: theory, Research and clinical applications*. New York: Guilford Press.

Cervone, D. & Shoda, Y. (1999). Social-cognitive theories and the coherence of personality. In D. Cervone & Y. Shoda (Eds.), *The coherence of personality* (pp. 3-36). New York: Guilford.

Chaiken, S., & Trope, Y. (Eds.). (1999). *Dual process theories in social psychology*. New York: Guilford.

Chen, S., & Chaiken, S. (1999). The heuristic-systematic model in its broader context. In S. Chaiken & Y. Trope (Eds.), *Dual process theories in social psychology* (pp. 73–96). New York: Guilford.

Ciancolo, AT., Matthew, C., Sternberg, R.J., & Wagner, R.K. (2006). Tacit knowledge, practical intelligence and expertise. In K.A. Ericsson, N. Charness, P.J., Feltovich, & R.R. Hoffman (Eds.), *The Cambridge handbook of expertise and expert performance* (pp. 613–632). New York: Cambridge University Press.

Cramer, P. (2006). *Protecting the self: Defense mechanisms in action*. New York: Guilford Press.

Damasio, A. (1994). *Descartes' error*. New York: Putnam.

Damasio, A. (1999). *The Feeling of What Happens*. New York: Harcourt and Brace.

de Waal, F. (1982). *Chimpanzee politics: Power and sex among apes*. London: Jonathan Cape.

de Waal, F. (1996). *Good natured: The origins of right and wrong in humans and other animals*. Cambridge, MA: Harvard University Press.

DeVries, R., & Zan, B. (1994). *Moral classrooms, moral children: Creating a constructivist atmosphere in early education*. New York: Teachers College Press.

Dijksterhuis, A. (2010). Automaticity and the unconscious. In S. Fiske, D. Gilbert, & G. Lindzey (Eds.), *Handbook of social psychology* (5th ed.) (Vol. I., pp. 228-67). New York: McGraw-Hill.

Ditto, P. H., Pizarro, D. A., & Tannenbaum, D. (2009). Motivated moral reasoning. In D. M. Bartels, C. W. Bauman, L. J. Skitka & D. L. Medin (Eds.), *The Psychology of Learning and Motivation* (Vol. 50, pp. 307–338). Burlington, VT: Academic Press.

Dupre, J. (2001). *Human Nature and the Limits of Science.* Oxford: Oxford University Press.

Feshbach, N.D. (1989). Empathy training and prosocial behavior. In J. Grobel & R.A. Hinde (Eds.), *Aggression and war: Their biological and social bases* (pp. 101–111). Cambridge: Cambridge University Press.

Fiske, A. P. (1991). *Structures of social life*. New York: Free Press.

Foot, P. (2001). *Natural Goodness*. Oxford: Clarendon Press.

Frankish, K. (2009). Systems and levels. In J. Evans and K. Frankish (Eds.), *Dual-system theories and the personal subpersonal distinction* (pp. 89-107). Oxford: Oxford University Press.

Frensch, P. A. (1998). One concept, multiple meanings: On how to define the concept of implicit learning. In M. A. Stadler and P. A. Frensch (Eds.) *Handbook of Implicit Learning* (pp. 47–104). Thousand Oaks, CA: Sage.

Fry, D. P., & Souillac, G. (2013). The relevance of nomadic forager studies to moral foundations theory: moral education and global ethics in the twenty-first century. *Journal of Moral Education*, 42(3), 346–359.

Gardner, S. (2000). Psychoanalysis and the personal/sub-personal distinction. *Philosophical Explorations*, 3(1), 96–119.

Gawronski, B. (2004). Theory-based bias correction in dispositional inference: the fundamental attribution error is dead, long live the correspondence bias. *European Review of Social Psychology*, 15, 183-217.

Gazzaniga, M. S. (1985). *The social brain.* New York: Basic Books.

Glaser, J. and Kihlstrom, J.F. (2005). Compensatory automaticity: Unconscious volition is not an oxymoron. In: Hassin, R.R., Uleman J.S., & Bargh J.A. (Eds). *The New Unconscious* (pp. 171–95). Oxford: Oxford University Press.

Goldie, P. (2007). Seeing What is the Kind Thing to Do. *Dialectica*, 61(3), 347–361.

Gollwitzer, P.M., Bayer, U.C., & McCulloch, K.C. (2005). The control of the unwanted. In: Hassin, R.R., Uleman J.S., & Bargh J.A. (Eds). *The New Unconscious* (pp. 485-515). Oxford: Oxford University Press.

Graham, J., Haidt, J., Koleva, S., Motyl, M., Iyer, R., Wojcik, S. P., & Ditto, P. H. (2013). Moral Foundations Theory: The Pragmatic Validity of Moral Pluralism. In Devine, P. & Plant, A. (Eds). *Advances in Experimental Social Psychology*, Vol. 47 (pp. 55–130). Burlington: Academic Press.

Greene, J.D., Sommerville, R.B., Nystrom, L.E., Darley, J.M., & Cohen, J.D. (2001). An fMRI study of emotional engagement in moral judgment. *Science*, 293, 2105–2108.

Greenspan, S. I., & Shanker, S. I. (2004). *The first idea.* Cambridge, MA: Da Capo Press.

Haidt, J. (2001). The emotional dog and its rational tail: A social intuitionist approach to moral judgment. *Psychological Review*, 108(4), 814–834.

Haidt, J. (2006). *The Happiness Hypothesis*. London: Arrow Books.

Haidt, J. (2010). Moral Psychology Must Not Be Based on Faith and Hope: Commentary on Narvaez (2010). *Perspectives on Psychological Science*, 5(2), 182–184.

Haidt, J. (2012). *The Righteous Mind.* London: Penguin.

Haidt, J., & Bjorklund, F. (2008). Social intuitionists answer six questions about moral psychology. In W. A. Sinnott-Armstrong (Ed.), *Moral Psychology* (Vol. 2, pp. 181–217). Cambridge, MA: MIT Press.

Haidt, J., & Hersh, M. A. (2001). Sexual morality: The cultures and reasons of liberals and conservatives. *Journal of Applied Social Psychology*, 31, 191–221.

Haidt, J., & Joseph, C. (2004, Fall). Intuitive ethics: How innately prepared intuitions generate culturally variable virtues. *Daedalus,* 55–66.

Haidt, J., & Kesebir, S. (2010). Morality. In S. Fiske, D. Gilbert, & G. Lindzey (Eds.), *Handbook of Social Psychology* (5th ed) (pp. 797–832). New York: McGraw-Hill.

Haidt, J., Graham, J., & Joseph, C. (2009). Above and Below Left–Right: Ideological Narratives and Moral Foundations. *Psychological Inquiry*, 20(2-3), 110–119.

Hamlin, J. K., Wynn, K., & Bloom, P. (2007). Social evaluation by preverbal infants. *Nature*, 450, 557–560.

Hammond, K. R. (2000). *Judgments Under Stress*. New York: Oxford University Press.

Harman, G., Mason, K. & Sinnott-Armstrong, W. (2010) Moral Reasoning. In J. Doris et al (Eds), *The Moral Psychology Handbook*. New York: Oxford University Press.

Hassin, R.R., Uleman J.S., & Bargh J.A. (Eds). *The New Unconscious*. Oxford: Oxford University Press.

Hogarth, R. M. (2001). *Educating Intuition*. Chicago: University of Chicago Press.

Huebner, B. (2011). Critiquing empirical moral psychology. *Philosophy of the Social Sciences*, 41(1), 50–83.

Hursthouse, R. (1999). *On Virtue Ethics*. Oxford: Oxford University Press.

Jacobson, D. (2005). Seeing by Feeling: Virtues, Skills, and Moral Perception. *Ethical Theory and Moral Practice*, 8, 387–409.

Juarrero, A. (1999). *Dynamics in Action: Intentional Behavior as a Complex System*. Cambridge, MA: MIT Press.

Kahneman, D., & Klein, G. (2009). Conditions for intuitive expertise. *American Psychologist*, 64, 515–526.

Keil, F.C., & Wilson, R.A. (2000). Explaining explanations. In F.C. Keil & R.A. Wilson (Eds.), *Explanation and cognition* (pp. 1–18). Cambridge, MA: MIT Press.

Kennett, J., & Fine, C. (2009). Will the Real Moral Judgment Please Stand Up?. *Ethical Theory and Moral Practice*, 12(1), 77–96.

Kiehl, K. A. (2006). A cognitive neuroscience perspective on psychopathy: Evidence for paralimbic system dysfunction. *Psychiatry Research*, 142, 107–128.

Klinger, E. (1978). Modes of normal conscious flow. In K.S. Pope & J.L. Singer (Eds.), *The stream of consciousness: Scientific investigations into the flow of human experience* (pp. 225–258). New York: Plenum.

Kochanska, G. (1997). Multiple pathways to conscience for children with different temperaments: From toddlerhood to age 5. *Developmental Psychology*, 33, 228-240.

Koenigs, M., Young, L., Adolphs, R., Tranel, D., Cushman, F., Hauser, M., & Damasio, A. (2007). Damage to the prefrontal cortex increases utilitarian moral judgments. *Nature*, 446, 865–866.

Kuhn, D. (1991). *The skills of argument*. Cambridge: Cambridge University Press.

Kunda, Z. (1990). The case for motivated reasoning. *Psychological Bulletin, JOS,* 480-498.

Kunda, Z. and Spencer, S.J. (2003). When do stereotypes come to mind and when do they color judgment? A goal-based theoretical framework for stereotype activation and application. *Psychological Bulletin*, 129, 522–544.

Lacewing, M. (forthcoming). Emotion and the virtues of self-understanding. In S. Roeser & C. Todd (Eds.) *Emotion and Value*. Oxford: Oxford University Press.

Lapsley, D., & Narvaez, D. (2004). A social-cognitive approach to the moral personality. In D. Lapsley & D. Narvaez (Eds.), *Moral development, self and identity* (pp. 189–212). Mahwah, NJ: Erlbaum.

Lazarus, R. S. (1991). Cognition and motivation in emotion. *American Psychologist*, 46, 352–367.

Lerner, J. S., & Tetlock, P. E. (2003). Bridging individual, interpersonal, and institutional approaches to judgment and decision making: The impact of accountability on cognitive bias. In S. L. Schneider & J. Shanteau (Eds.), *Emerging perspectives on judgment and decision research* (pp. 431–457). New York: Cambridge.

Lerner, J.S., Goldberg, J.H., & Tetlock, P.E. (1998). Sober second thought: the effects of accountability, anger, and authoritarianism on attributions of responsibility. *Personality and Social Psychology Bulletin*, 24, 563–574.

Lewis, M.D. (2009). Desire, dopamine, and conceptual development. In S.D. Calkins & M.A. Bell (Eds.), *Child development at the intersection of emotion and cognition* (pp. 175–199) Washington, DC: American Psychological Association.

Lilienfeld, S.O., Ammirati, R., & Landfield, K. (2009). Giving debiasing away: Can psychological research on correcting cognitive errors promote human welfare? *Perspectives on Psychological Science*, 4, 390–398

MacLean, P. D. (1990). *The triune brain in evolution.* New York: Plenum.

McAdams, D. P. (1995). What do we know when we know a person? *Journal of Personality*, 63(3), 365–396.

McDowell, J. (1979). Virtue and Reason. *The Monist*, 62, 331–50.

McDowell, J. (1985). Values and Secondary Qualities. In T. Honderich (Ed.), *Morality and Objectivity: a Tribute to J.L. Mackie* (pp. 110–129), London: Routledge.

Mercier, H. and Sperber, D. (2011). Why do humans reason? Arguments for an argumentative theory. *Behavioral and Brain Sciences*, 34, 57–111.

Mischel, W. (1968). *Personality and assessment*. Hoboken, NJ: John Wiley & Sons.

Mischel, W. & Shoda Y. (1995). A cognitive-affective system theory of personality: reconceptualizing situations, dispositions, dynamics, and invariance in personality structure. *Psychological Review*, 102, 246–68.

Moll, J., de Oliveira-Souza, R., & Eslinger, P.J. (2003). Morals and the human brain: A working model. *Neuroreport*, 14, 299-305.

Moll, J., de Oliveira-Souza, R., Bramati, I., & Grafman, J. (2002). Functional networks in emotional moral and nonmoral social judgments. *NeuroImage*, 16, 696-703.

Monteith, M.J., Ashburn-Nardo, L., Voils, C.I., and Czopp, A.M. (2002). Putting the brakes on prejudice: on the development and operation of cues for control. *Journal of Personality and Social Psychology*, 83, 1029–1050.

Moskowitz, G. B., Skurnik, L, & Galinsky, A. D. (1999). The history of dual process notions, and the future of pre-conscious control. In S. Chaiken & Y. Trope (Eds.), *Dual process theories in social psychology* (pp. 12-36). New York: Guilford Press.

Narvaez, D. (1998). The influence of moral schemas on the reconstruction in eighth-graders and college students. *Journal of Educational Psychology*, 90, 13-24.

Narvaez, D. (2005). Integrative ethical education. In M. Killen & J. Smetana (Eds.), *Handbook of moral development* (pp. 703–733). Mahwah, NJ: Erlbaum.

Narvaez, D. (2008a). The Social Intuitionist Model: Some Counter-Intuitions. In W. A. Sinnott-Armstrong (Ed.), *Moral Psychology* (Vol. 2, pp. 233–240). Cambridge, MA: MIT Press.

Narvaez, D. (2008b). Triune ethics: The neurobiological roots of our multiple moralities. *New Ideas in Psychology*, 26, 95–119.

Narvaez, D. (2010). Moral Complexity The Fatal Attraction of Truthiness and the Importance of Mature Moral Functioning. *Perspectives on Psychological Science*, 5(2), 163–181.

Narvaez, D. (2013a). Wisdom as Mature Moral Functioning: Insights from Developmental Psychology and Neurobiology. In M. Jones, P. Lewis, and K. Reffitt (Eds.), *Toward Human Flourishing: Character, Practical Wisdom and Professional Formation*. Macon, GA: Mercer University Press.

Narvaez, D. (2013b). The 99 Percent—Development and Socialization Within an Evolutionary Context. In D. Fry (Ed.), *War, Peace and Human Nature* (pp. 341–357). New York: Oxford University Press.

Narvaez, D. (2013c). Neurobiology and moral mindset. In K. Heinrichs, F. Oser & T. Lovat (Eds.), *Handbook of Moral Motivation* (pp. 323–340). Rotterdam: Sense Publishers.

Narvaez, D., & Bock, T. (in press). The development of virtue. In L. Nucci & D. Narvaez (Eds.), *Handbook of Moral and Character Education* (2nd ed.). New York: Routledge.

Narvaez, D., & Lapsley, D. K. (2005). The Psychological Foundations of Everyday Morality and Moral Expertise. In D. Lapsley & Power, C. (Eds.), *Character psychology and character education* (pp.140–165). Notre Dame: IN: University of Notre Dame Press.

Narvaez, D., Getz, I., Rest, J.R., & Thoma, S. (1999). Individual moral judgment and cultural ideologies. *Developmental Psychology*, 35, 478–488.

Narvaez, D., Lapsley, D.K., Hagele, S., & Lasky, B. (2006). Moral chronicity and social information processing: Tests of a social cognitive approach to the moral personality. *Journal of Research in Personality,* 40, 966–985.

Nelson, E.E., & Panksepp, J. (1998). Brain substrates of infant–mother attachment: contributions of opioids, oxytocin, and norepinephrine. *Neuroscience and Biobehavioral Reviews*, 22, 437–452.

Nickerson, R. S. (1994). The teaching of thinking and problem solving. In R. J. Sternberg (Ed.), *Thinking and problem solving* (pp. 409-449). San Diego: Academic Press.

Nickerson, R.S. (1998). Confirmation bias: a ubiquitous phenomenon in many guises. *Review of General Psychology,* 2(2), 175-220.

Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, 84, 231-259.

Nussbaum, M. (1993). Non-relative virtue. In M. Nussbaum & A. Sen (Eds.), *The Quality of Life* (pp. 242–269). Oxford: Clarendon Press.

Panksepp, J. (1998). *Affective neuroscience.* New York: Oxford University Press.

Panksepp, J., & Panksepp, J.B. (2000). The seven sins of evolutionary psychology. *Evolution and Cognition*, 6(2), 108-131.

Pascarella, E., & Terenzini, P. (1991) *How college affects students.* San Francisco: Jossey-Bass.

Payne, B., Jacoby, L.L., & Lambert, A.J. (2005). Attitudes as accessibility bias: Dissociating automatic and controlled processes. In R.R. Hassin, J.S. Uleman, & J.A. Bargh (Eds). *The New Unconscious* (pp. 393–420). Oxford: Oxford University Press.

Perkins, D.N., Farady, M., & Bushey, B. (1991). Everyday reasoning and the roots of intelligence. In J.F. Voss, D.N. Perkins, & J.W. Segal (Eds.), *Informal reasoning and education* (pp. 83–105). Hillsdale, NJ: Erlbaum.

Pettigrew, T.F., & Tropp, L.R. (2000). Does intergroup contact reduce prejudice? Recent meta-analytic findings. In S. Oskamp (Ed.), *Reducing prejudice and discrimination* (pp. 93–114). Mahwah, NJ: Erlbaum.

Pizarro, D.A., & Bloom, P. (2003). The intelligence of the moral intuitions: Comments on Haidt (2001). *Psychological Review*, 110, 193–196.

Plous, S. (2003). The psychology of prejudice, stereotyping, and discrimination: An overview. In S. Plous (Ed.), *Understanding prejudice and discrimination* (pp. 3–48). New York: McGraw-Hill.

Power, F.C., Higgins, A., & Kohlberg, L. (1989). *Kohlberg's approach to moral education.* New York: Columbia University Press.

Prinz, J. (2004). *Gut reactions*. New York: Oxford University Press.

Prinz, J. (2007). *The emotional construction of morals*. New York: Oxford University Press.

Railton, P. (forthcoming). The affective dog and its rational tale. *Ethics.*

Rest, J.R. (1979). *Development in judging moral issues.* Minneapolis: University of Minnesota Press.

Sauer, H. (2011). Social intuitionism and the psychology of moral reasoning. *Philosophy Compass*, 6(10), 708–21.

Sauer, H. (2012). Educated intuitions. Automaticity and rationality in moral judgement. *Philosophical Explorations*, 15(3), 255–275.

Schnall, S., Haidt, J., Clore, G. L., & Jordan, A. H. (2008). Disgust as embodied moral judgment. *Personality and Social Psychology Bulletin*, 34, 1096–1109.

Schore, A. (2001). The effects of a secure attachment on right brain development, affect regulation and infant mental health. *Infant Mental Health Journal*, 22, 201–269.
Schore, A. (2003). *Affect dysregulation and disorders of the self*. New York: Norton.
Selman, R.L. (2003). *The promotion of social awareness: Powerful lessons from the partnership of developmental theory and classroom practice.* New York: Russell Sage.
Snow, N.E. (2010). *Virtue as social intelligence: An empirically grounded theory*. New York: Routledge.
Sperber, D. (2005). Modularity and relevance: How can a massively modular mind be flexible and context-sensitive? In P. Carruthers, S. Laurence & S. Stich (Eds.), *The innate mind: Structure and contents* (pp. 53-68). New York: Oxford.
Stern, D. N. (1985). *The interpersonal world of the human infant*. New York: Basic Books.
Taylor, S. & Brown, J. (1988). Illusion and Well-Being: A Social Psychological Perspective on Mental Health. *Psychological Bulletin*, 103, 193–210.
Taylor, S. & Brown, J. (1994). Positive Illusions and Well-Being Revisited: Separating Fact from Fiction. *Psychological Bulletin*, 116, 21–27.
Thoma, S.J., Barnett, R., Rest, J., & Narvaez, D. (1999). Political identity and moral judgment development using the Defining Issues Test. *British Journal of Social Psychology*, 38, 103–111.
Tiberius, V. (2008). *The Reflective Life*. Oxford: Oxford University Press.
Tooby, J., Cosmides, L., & Barrett, H. C. (2005). Resolving the debate on innate ideas: Learnability constraints and the evolved interpenetration of motivational and conceptual functions. In P. Carruthers, S. Laurence, & S. Stich (Eds.), *The innate mind: Structure and contents* (pp. 305–337). New York: Oxford University Press.
Trope, Y. and Fishbach, A. (2005). Going beyond the motivation given: self-control and situational control over behaviour. In R.R. Hassin, J.S. Uleman, & J.A. Bargh (Eds). *The New Unconscious* (pp. 537-565). Oxford: Oxford University Press.
Uleman, J.S., Winborne, W.C., Winter, L., & Schechter, D. (1986). Personality differences in spontaneous trait inferences at encoding. *Journal of Personality and Social Psychology*, 51, 396-404.
Vaillant, G. E. (1993). *The Wisdom of the Ego*. Cambridge, MA: Harvard University Press.
Vaillant, G. E. (2000). Defense mechanisms. In A. Kazdin (Ed.), *Encyclopedia of psychology* (vol. 2, pp. 454-457). Washington, DC: American Psychological Press, 2000.
Varela, F. (1999). *Ethical know-how.* Stanford, CA: Stanford University Press.
Warneken, F., & Tomasello, M. (2006). Altruistic helping in human infants and young chimpanzees. *Science,* 311, 1301–1303.
Wegner, D. and Bargh, J. (1998). Control and automaticity in social life. In S. Fiske, D. Gilbert, & G. Lindzey (Eds), *Handbook of social psychology* (4th ed.) (Vol. I., pp. 446-96). New York: McGraw-Hill.
Westen, D. (2007). *The political brain.* New York: Public Affairs.
Wheatley, T., & Haidt, J. (2005). Hypnotic disgust makes moral judgments more severe. *Psychological Science*, 16, 780–784.
Wiggins, D. (1987). *Needs, values, truth*. Oxford: Blackwell.
Willingham, D.T. (2007). Critical thinking: Why is it so hard to teach? *American Educator*, 31, 8–19.
Wilson, T. D. (2002). *Strangers to Ourselves: Discovering the Adaptive Unconscious*. Cambridge, MA: Harvard University Press.
Wright, J. C. & Baril, G. L. (2013). Understanding the role of dispositional and situational threat sensitivity in our moral judgments. *Journal of Moral Education*, 42 (3), 383-397.
Zelli, A., Huesmann, L.R., & Cervone, D. (1995). Social inference and individual differences in aggression: Evidence for spontaneous judgements of hostility. *Aggressive Behaviour*, 21, 405-417.