

# Moral relativism and evolutionary psychology

Steven D. Hales

Received: 1 May 2007 / Accepted: 5 May 2007 / Published online: 21 December 2007  
© Springer Science+Business Media B.V. 2007

**Abstract** I argue that evolutionary strategies of kin selection and game-theoretic reciprocity are apt to generate agent-centered and agent-neutral moral intuitions, respectively. Such intuitions are the building blocks of moral theories, resulting in a fundamental schism between agent-centered theories on the one hand and agent-neutral theories on the other. An agent-neutral moral theory is one according to which everyone has the same duties and moral aims, no matter what their personal interests or interpersonal relationships. Agent-centered moral theories deny this and include at least some prescriptions that include ineliminable indexicals. I argue that there are no rational means of bridging the gap between the two types of theories; nevertheless this does not necessitate skepticism about the moral—we might instead opt for an ethical relativism in which the truth of moral statements is relativized to the perspective of moral theories on either side of the schism. Such a relativism does not mean that any ethical theory is as good as any other; some cannot be held in reflective equilibrium, and even among those that can, there may well be pragmatic reasons that motivate the selection of one theory over another. But if no sort of relativism is deemed acceptable, then it is hard to avoid moral skepticism.

**Keywords** Moral relativism · Agent-centered · Agent-neutral · Evolutionary psychology · Intuition

## 1 Introduction

Evolutionary psychological explanations of ethical norms tend to be darkly pessimistic about the truth of any moral propositions, doubting that instincts and feelings formed

---

S. D. Hales (✉)  
Department of Philosophy, Bloomsburg University, Bloomsburg, PA 17815, USA  
e-mail: hales@bloomu.edu

over millions of years by blind impersonal forces connect us at all with an abstract realm of moral reality. Either that or they are cheerily optimistic about locating in all nondefective human beings, from every culture and epoch, a universal sense of right and wrong that is our inheritance from Mother Nature. In this article I sketch out something of a *via media* between the optimists and the pessimists. I concur that evolutionary psychology offers the best hope of an adequate explanation of moral intuition. I then argue that evolutionary strategies of kin selection and game-theoretic reciprocity are apt to generate agent-centered and agent-neutral intuitions, respectively. Such intuitions are the building blocks of moral theories, resulting in a fundamental schism between agent-centered theories on the one hand and agent-neutral theories on the other. An agent-neutral moral theory is one according to which everyone has the same duties and moral aims, no matter what their personal interests or interpersonal relationships. Agent-centered moral theories deny this and include at least some prescriptions that include ineliminable indexicals. I argue that there are no rational means of bridging the gap between the two types of theories; nevertheless this does not necessitate skepticism about the moral—we might instead opt for an ethical relativism in which the truth of moral statements is relativized to the perspective of moral theories on either side of the schism. Such a relativism does not mean that any ethical theory is as good as any other; some cannot be held in reflective equilibrium, and even among those that can, there may well be pragmatic reasons that motivate the selection of one theory over another. But if no sort of relativism is deemed acceptable, then it is hard to avoid moral skepticism.

## 2 Moral epistemology

The familiar model of moral epistemology is that of reflective equilibrium.<sup>1</sup> According to this model, the development of moral theories proceeds through the use of rational intuition that generates immediate, spontaneous judgments about specific cases of potential behavior. These intuitions are then used as data to develop more sophisticated theories through reflective equilibrium. In a procedure akin to the development of scientific theories, our spontaneous moral judgments about what acts or types of acts to perform in specific cases are treated as raw data. If there is no inherently plausible theory that perfectly matches our intuitions (i.e., can incorporate all of the raw data), then what we do is develop the best theory that we can to accommodate as many of our preanalytic intuitions as possible. The remainder of the intuitive data are treated as outliers or it is assumed that there are other undiscovered factors that will eventually incorporate that data into the theory. Thus it is a matter of treating our intuitions as *prima facie*, but not sacrosanct, data and then building and revising our theories to best fit them. I have argued elsewhere that this methodology is not unique to the development of normative theories, but is in fact the usual procedure in all of analytic rationalist philosophy.<sup>2</sup>

<sup>1</sup> See McMahan (2000, p. 100), “The most commonly endorsed method of moral inquiry among contemporary moral philosophers is the method described by John Rawls under the label ‘reflective equilibrium’.”

<sup>2</sup> Hales (2006).

One of the striking features about contemporary ethics is the persistent threat of the plurality problem. This is the problem of mutually incompatible theories that are equally internally coherent, consistent, capable of accommodating the preanalytic data, and with component parts that are explanatorily connected. Consider the apparent standoff between consequentialist and deontological theories. No outside third party could reasonably deny that both theoretical approaches accommodate many, but not all, of our ordinary moral intuitions, that both are well-developed, detailed, internally coherent theories and that yet there are cases in which the theories produce mutually inconsistent moral instruction. It is a seeming Mexican standoff, with partisans firing ever more baroque counterexamples at each other while the defenders dodge, ruling those cases to be outliers. John Mackie took the plurality problem as a reason to reject moral realism and embrace a form of skepticism.<sup>3</sup> The main rejoinder to Mackie is to insist that moral disagreement is not intractable at all, and to stay the course with a troop surge and better weaponry.

I concur with Mackie that the plurality problem in ethics is indeed intractable, and further that the lesson to draw is that we should be either moral relativists or moral skeptics. My argument for this conclusion is based upon a more detailed examination of the nature of moral intuition. Jaakko Hintikka has complained that “blind faith [in intuition] is below the intellectual dignity of philosophers for whom unexamined intuitions should not be worth intuiting” (Hintikka 1999, p. 130). I will attempt to answer Hintikka’s challenge in the case of moral intuitions,<sup>4</sup> suggesting that recent work in evolutionary psychology offers the most promising means of explaining why we have moral intuitions at all and why we have the ones that we do. I will not offer a general defense of the evolutionary psychology approach to explaining morality, and a die-hard rationalist who thinks that ethics is *sui generis* and emancipated from the natural world will find most of what I have to say unpersuasive. Rather, I will argue that once we look at even a few of the details of how natural selection has shaped, or even might have shaped, our moral instincts, the hope that unlimited reasoning about cases will eventually yield agreement about moral theory begins to fade. Fundamental dichotomies in our moral intuitions have their roots in distinct evolutionary strategies, and it is these dichotomies that are the respective bases for the chief varieties of moral theory. In particular, I believe that agent-centered and agent-neutral intuitions are underwritten by different evolutionary tactics.

### 3 EP explanations of morality

There is an enormous amount of recent work on evolutionary explanations of moral phenomena, with several books and articles on the topic appearing last year alone. While I lack both the space and the will to present a comprehensive defense of evolutionary psychology in this article, it is worth noting that evolutionary psychology (EP) has numerous virtues as an explanation of moral phenomena. First, it serves to explain why there are widespread values universal across all known human cultures

<sup>3</sup> Mackie (1977, Chap. 1).

<sup>4</sup> I discuss rational intuition more generally and in more detail in Hales (2006).

by demonstrating that innate to our species are certain sorts of moral imperatives and that these imperatives are part of basic evolutionary human design. Second, it serves to explain why there is enormous agreement between otherwise diverse moral theories. Third, it connects human morality to moral, pre-moral, or quasi-moral behavior in other species, such as altruism, deception and punishment, pair-bonding, helping, and caring, and has additional empirical support from brain studies of moral responses.<sup>5</sup> Fourth, EP provides a means of explanatory unity between what has been traditionally regarded as the pure a priori domain of moral theory and the natural and social sciences. These reasons may not be definitive, but no other account of moral intuition is nearly as promising or exhibits the same range of virtues.

There are a variety of possible evolutionary options. On the one extreme is strict genetic determinism: all intuitive responses to moral scenarios are hard-coded into our genes, and this hard coding is a component of the genetic heritage for *homo sapiens*. The genetic determinism option is very unlikely for multiple reasons. The first is that it is probably impossible that the human genome could encode for every contingency; there are just too many possible scenarios that require moral judgment.<sup>6</sup> The second reason is that there is widespread moral disagreement about cases—there is nothing like univocality about abortion, capital punishment, the treatment of nonhuman animals, sexual ethics, permissible wartime behavior, and the like. Strict genetic determinism cannot explain this divergence. There is no “morality gene,” and it is unlikely that any serious thinker has ever thought otherwise.

The other extreme is that there is no shared biological basis for moral intuitions. They are nothing more than the prejudices of our own societies, inculcated in us from the time we are young, now so ingrained that they have the force of truth. Russell once accused Kant of doing no more than regurgitating the morality that he unreflectively imbibed at his mother’s knee.<sup>7</sup> In this way perhaps our own moral intuitions are no more than our intuitions about valves—the faucets in my house open when turned counter-clockwise and close when turned clockwise. When I encounter an unknown valve in a hotel room my intuition and expectation is that it will work exactly the same way. Yet this is purely cultural, since we could have all subscribed to the convention to build valves that work the opposite way and our intuitions would be similarly reversed. This extreme is also very implausible, primarily because there is widespread, cross-cultural, and considerable intuitive agreement about a great many moral matters—that it is good to help others, that incest and indiscriminate killing are wrong, that one has familial duties, and so on. Of the 373 traits that anthropologist Donald E. Brown identifies as human universals, the majority involve moral or proto-moral judgment and behavior.<sup>8</sup> If morality is purely a cultural phenomenon, wholly divorced from the biological nature of our species, then there is no particular reason that societies diverse in time and place should have any intersecting moral precepts at all beyond whatever minimal constraints

<sup>5</sup> See in particular de Waal (2006) and Gazzaniga (2005).

<sup>6</sup> Compare Joyce (2006, pp. 6–7).

<sup>7</sup> Russell (1927, p. 590).

<sup>8</sup> See Brown (1991, Chap. 6). Shermer (2004, Appendix II), claims that 202 of Brown’s 373 universals are related to morality.

are needed for groups of humans to survive until reproduction. But they all do. A lack of a shared biological basis leaves moral universals unexplained, if not unexplainable.

The most likely evolutionary option is the mean: built-in common morality is very general, incorporating ingrained intuitive responses to a few specific types of cases while also allowing considerable environment-responsive plasticity. The responses to such cases will diverge rapidly along the varying pathways which our rootstock of moral intuitions can grow. As Hauser has recently argued with some vigor (Hauser 2006), it is analogous to holding that language is innate, but not Mandarin, Magyar, or Spanish.<sup>9</sup> One's linguistic intuitions will be language-specific, even though evolution has built-in considerable linguistic plasticity. Thus, as a native English speaker, I find it natural and intuitive to place adjectives before nouns, to conjugate my verbs but not decline my nouns, that nouns are ungendered, that prepositions always come before their noun phrases, and plenty of other things that do not hold in all languages. In many other languages, speakers find none of these things intuitive, even though there are super-rules, or deep structure, that governs all syntax in all languages.<sup>10</sup>

Hauser has argued that just as our evolution-built deep syntax guides or regulates all human languages, so too there is a universal moral grammar that underwrites all specific moral systems. If he is right about this, then we should expect to find somewhat vague universals holding in every human society, but varying in interpretation. Thus, proscriptions like “it is good to help others” or “incest is wrong” are everywhere, but their application is not universal— help which others? all others? my family and friends?, incest is wrong, but who counts as kin? Uncles, aunts, and cousins, or not?. That is to say, the moral proscription against incest (for example) built by natural selection will manifest itself as intuitions concerning the repugnancy of sexual congress with specific groups of individuals. Which groups count as kin and therefore sexually off-limits will vary by culture. In fact, we find exactly the expected variation not only across different cultures, but also across well-developed moral theories.

One might suspect at this point that a “universal moral grammar” is not going to give rise to any particular moral theory, that deontology, utilitarianism, or any number of other theories could arise out of it and be compatible with whatever the deep syntax of our morality is. Sticking with Hauser's basic analogy, there is no more reason to think that one unique moral theory will be the result of our basic moral coding than there is reason to think that the existence of Chomskian deep grammar will result in just one human language. Acceptance of the evolutionary origins of varying moral codes is not yet an adequate reason to be a moral skeptic or moral relativist, however. It could be that pure reason and rational argument can guide us to the one true moral theory, no matter what the biologically distasteful origins of those theories. While one might view such a project as akin to the Carnapian quest for a logically perfect language, a satisfactory argument for moral relativism will have to show that the fundamental intuitions that are used as data for the development of moral theories are themselves an inconsistent conglomeration produced by varying evolutionary strategies.

<sup>9</sup> de Waal (2006, pp. 166–167) also endorses the language analogy.

<sup>10</sup> This is the Chomsky point, nicely explained in Pinker (1994, Chap. 4).

## 4 Two evolutionary strategies for building moral intuitions

In his excellent recent book on these matters, Richard Joyce argues that there are four primary evolutionary forces that underwrite helping behaviors, behaviors that in turn are the fundament for fitness-sacrificing altruism and moral judgment. These are kin selection, group selection, direct reciprocity and indirect reciprocity. Not everyone agrees that these are distinct strategies or that they are of equal importance in understanding the evolution of morality. Sober and Wilson (1998) regard kin selection as a special case of group selection, and Dawkins (2006) considers group selection to be possible but rare. Fortunately, settling these debates is unnecessary for the points I wish to make presently. What is commonly agreed upon—and this is key—is that these are the evolutionary forces that ultimately produce moral intuitions.<sup>11</sup>

What I will argue for is this: the intuitions we should expect from kin selection are agent centered ones, whereas the intuitions we should expect from reciprocity are agent neutral.<sup>12</sup> The preceding thesis is, as far as I know, undefended so far in the literature.

### 4.1 Kin selection

There is considerable empirical evidence that the more genetically similar organisms are, the more they are liable to help each other even at a personal cost. Hume presciently wrote that

A man naturally loves his children better than his nephews, his nephews better than his cousins, his cousins better than strangers, where everything else is equal. Hence arise our common measures of duty, in preferring the one to the other. Our sense of duty always follows the common and natural course of our passions. (Hume 1978 (1739)) bk. III, pt. 2, Sect. 1.

According to Hamilton's Rule (1964), a trait of helping others at some cost to the individual can be expected to be favored by natural selection if  $rB > C$ , where  $r$  is the degree of genetic relatedness to the individual,  $B$  is the benefit to the recipient, and  $C$  is the cost to the individual. Taken alone, it is thus puzzling how helping behaviors, much less moral intuitions, directed towards non-kin might arise. If  $r=0$ , then it doesn't matter how much benefit non-kin would receive by helping; the cost to the putative helper is always going to be too high (since the cost will always be  $>0$ ). There may be other evolutionary forces at work (Hamilton's Rule does not state that we should expect to find a trait of helping others *only if*  $rB > C$ ) but there is a *prima facie* case that kin selection won't do the trick.

<sup>11</sup> Wright (1994, Chap. 16), Ridley (1996, Chap. 7), Wilson (1998, Chap. 11), Flack and de Waal (2000), Shermer (2004, Chap. 2), Hauser (2006), de Waal (2006), and Lieberman (2007). For some reason, much of the discussion in this area uses the 18th century terminology "moral sentiments" instead of the more modern "moral intuitions."

<sup>12</sup> Axelrod (1984, p. 89) and de Waal (2006, p. 16) explicitly characterize kinship theory and reciprocity theory as the two chief ways to give an evolutionary account of cooperation. De Waal, p. 53, goes so far as to characterize kin selection and reciprocity as the "foundations" of an evolutionary account of morality.

The prima facie case is not dispositive, though. It turns out that kin selection *is* able to provide an explanation for the fact that people are more likely to help non-kin with whom they interact socially—or even have greater spatial proximity to—than distant strangers who are non-kin. What we must consider is the mechanism used by natural selection to segregate kin from non-kin. While we might use scent, or possibly phenotypic markers, to make the distinction as do other species, these strategies do not seem to be the ones used by humans. Lieberman et al. (2003) have recently argued that the innate human kin-recognition system is approximately this: kin are those with whom you are familiar from childhood. The evidence they provide involves the incest taboo. Childhood co-residence with opposite-sex individuals is a strong and accurate predictor of the strength of moral intuitions (or sentiments) regarding opposition to incest. Co-residents who knew that they were not genetically related felt little sexual attraction towards each other, even in societies where they had been co-reared as future spouses. Moreover, co-residence duration alone predicted the presence and strength of moral intuitions opposing 3rd party sibling incest, even after controlling for several other variables.<sup>13</sup>

Joyce suggests analogously that the helping mechanism generated by kin selection is something like “provide-help-to-those-conspecifics-with-whom-you-interact-frequently” Joyce (2006, p. 22). Genetic relatedness is not readily observable, but co-residence or frequent interaction is; therefore it is easier for natural selection to utilize the latter as a stalking-horse for the former. If kin selection is thereby one of the underlying mechanisms for the production of desires or instincts to help those with whom one interacts, most often family and friends, then it is a short step to conclude that it is the basis of moral intuitions to aid one’s family or friends. As Darwin wrote in *The Descent of Man*,

It must not be forgotten that although a high standard of morality gives but a slight or no advantage to each individual man and his children over the other men of the same tribe, yet that an increase in the number of well-endowed men and an advancement in the standard of morality will certainly give an immense advantage to one tribe over another. A tribe including many members who, from possessing in a high degree the spirit of patriotism, fidelity, obedience, courage, and sympathy, were always ready to aid one another, and to sacrifice themselves for the common good, would be victorious over most other tribes; and this would be natural selection. At all times throughout the world tribes have supplanted other tribes; and as morality is one important element in their success, the standard of morality and the number of well-endowed men will thus everywhere tend to rise and increase. Darwin (1882, p. 132).

Joyce, however, argues that kin selection “could hardly explain human *morality*, in which (in at least the Western tradition) the tendency to favor one’s own family members is a vice to which we have given the name ‘nepotism’” Joyce (2006, p. 21). Unfortunately, Joyce is too hasty here for two different reasons. The first reason is that an evolutionary psychology account of morality cannot assume that *Western* moral norms

<sup>13</sup> Lieberman et al. (2003). Compare Wilson (1998, Chap. 7).

represent “true” morality. An adequate evolutionary explanation of morality must be able to encompass all strands of moral thinking, intuitions, rules, and judgments. In many non-Western societies tribalism and kinship have an extensive and serious importance for establishing moral obligations and commitments. For example, the Pashtuns of Afghanistan and Pakistan have a greater sense of loyalty and duty towards their fellow tribesmen than to the nation-states in which they reside, and this is not regarded by their fellow citizens as unexpected or mistaken in the least. Indeed, a failure to appreciate the strong sense of moral obligation associated with tribes helps to explain the ineffective political approach of the United States in the Middle East, which seems to assume that “liberated” peoples will naturally choose agent-neutral democracy.<sup>14</sup>

Second, and most importantly for the present essay, the idea that one has special duties to one’s own family and friends is the centerpiece of agent-centered or agent-relative morality. An agent-centered moral theory is one that includes at least some obligations that are essentially indexical—some of the things that the theory tells me I ought to do conflict with the things that the things the theory says that you ought to do. Agent-centered intuitions are ones prompted by cases such as these. *Drowning Swimmers*. You can save two drowning swimmers but not both. You are the only means of rescue. One of the drowning swimmers is your child and the other is a stranger child with equal life prospects and a comparable network of family and friends. While from the perspective of net utility it is morally indifferent which child you save, most people have a powerful intuition that they have a special duty to their own child. Not merely do most feel a preference to save one’s own child, but in addition a sense that it is one’s moral obligation to do so. *A Friend in Need*. Your best friend is down on her luck—her husband left her and she is looking for work. She is having trouble making ends meet and you are considering buying some groceries for her to tide her over. As you enter the grocery store, you see a charity collecting outside to send money to aid famine victims in a distant land you have never visited. The charity claims that you will save lives through the donation of a few dollars. Not being terribly flush yourself, you can either buy groceries for your best friend or donate to the charity, but not both. Even though the expected utility of donating to the charity is higher, most people feel more of a duty to help the friend.

What explains why most people have intuitions that their primary moral obligation is to save their own child or to help their troubled friend? My contention is that these intuitions are the result of evolutionary mechanisms like kin selection. If the kin selection imperative of “provide-help-to-those-conspecifics-with-whom-you-interact-frequently” is hardwired in, we should expect agent-centered intuitions to be the output. We instinctively help our families and friends, and feel that we *ought* to help them, precisely the sort of instincts built by kin selection.

However, those are not the only sort of moral intuitions we have. To explain how agent-neutral intuitions could arise out of natural selection, let us turn to

#### 4.2 Reciprocity

We are often in circumstances in which we might aid others, and the value of our aid exceeds the cost of providing it. These are precisely the circumstances in which

<sup>14</sup> A point that is occurring somewhat belatedly to popular commentators. See for example Lowry (2006).



consequentialists will insist that our moral duty is to provide such aid. But how could evolutionary theory provide an explanation as to why we would have intuitions to make even minor sacrifices for unrelated, unfamiliar individuals? The answer is that it is fitness-advancing for individuals to provide help to others if it is likely there will be reciprocal assistance from either those helped (direct reciprocity) or by those in the helper's group who consider the helper's status raised as a result of his or her generosity (indirect reciprocity). As Joyce writes,

If a male is helpful to a female (bringing her food, etc.) and as a result she confers on him the proportionally greater benefit of reproduction, this is an example of direct reciprocity. If a male is helpful to his fellows in general, and as a result an observant female confers on him the proportionally greater benefit of reproduction (thus producing sons who are generally helpful and daughters who have a preference for helpful males), this is an example of indirect reciprocity. Joyce (2006, p. 33).

There are several non-hypothetical examples of such behavior in among non-human animals.<sup>15</sup> More generally, the application of game-theoretic models (prominently iterated multi-player simultaneous choice games) to evolutionary design shows how natural selection would plump for organisms that are motivated to make certain sacrifices to aid others, even when there is no guarantee of reciprocal help, and even when the other players are unfamiliar non-kin.

Consider the much-discussed example of the Prisoner's Dilemma. Two-player Prisoner's Dilemmas are characterized by a specific payoff structure. If both players cooperate, both receive the reward payoff (R), if one cooperates and the other defects then the cooperator receives a sucker payoff (S) and the defector receives the temptation payoff (T), and finally if both players defect then both receive the punishment payoff (P). In the Prisoner's Dilemma,  $T > R > P > S$ . What makes this such a dilemma, of course, is that the Nash equilibrium point<sup>16</sup> is for both players to defect and both receive the punishment payoff. Each player fears that he will cooperate and the other will defect (yielding the temptation payoff for the defector and the sucker payoff for the cooperator), so they both defect. Mutual cooperation maximizes net utility, but it is difficult to see how actors could rationally opt for that solution when simultaneously choosing to cooperate or defect. In one-off games, defection seems almost unavoidable. However, in the case of iterated games, natural selection will favor mutual cooperators, and after sufficient evolutionary time, their descendants will instinctually cooperate.

A real life iterated Prisoner's Dilemma is in the case of the ziczac bird, which cleans parasites out of the mouths of crocodiles. This is a case of reciprocal aid: the ziczac helps the crocodile by removing parasites and the crocodile helps the ziczac by providing a meal. Both the ziczac and the crocodile could cheat, or defect, from the

<sup>15</sup> See, Trivers (1971) for classic cases. For a recent discussion of reciprocity in chimpanzees and capuchin monkeys, see de Waal (2006, p. 42) *et passim*.

<sup>16</sup> A pair of moves is a Nash equilibrium point when each move is the best reply to the other. Put another way, it is the solution such that after completion of the game both players are unable to Monday-morning quarterback a better choice than the one they actually made in the game.

arrangement—the ziczac might help himself to some tender crocodile cheek while he is in there eating parasites, and the crocodile could get both a cleaning and a free meal of ziczac bird. Here is a decision matrix that illustrates the nature of their predicament (the numbers are irrelevant so long as they satisfy the Prisoner’s Dilemma characteristic ordering of  $T > R > P > S$ ):

	Ziczac cooperates (cleans crocodile)	Ziczac defects (nips croc cheek)
Croc cooperates (allows cleaning)	Croc 2, Ziczac 2	Croc 0, Ziczac 3
Croc defects (eats ziczac)	Croc 3, Ziczac 0	Croc 1, Ziczac 1

If such a game were played only once, it would be deeply puzzling as to why a ziczac would chance a meal inside the mouth of a crocodile, or, if he did, why the crocodile would refrain from eating him. After all, ziczacs and crocodiles are unrelated species who have no bonds built up by other forms of interaction and have no kinship ties. But the game is not played just once—and it is through repeated iterations over a long period of time that natural selection can promote mutual cooperation, the solution of greatest overall utility. Just how this would work has been extensively addressed by game theorists and biologists.<sup>17</sup> In one sense, it should not seem terribly surprising; evolution is nothing if not patient, and the great glory of natural selection is the explanation of how things extremely improbable on their face can develop by tiny changes over long expanses of time. Reciprocity, unlike kin selection, is a way of building relationships with unknown individuals.

In pure iterated prisoner’s dilemmas, the players have no *expectations* that the game will be played more than once. In the case of animals like ziczacs and crocodiles, or other similar examples like cleaner fish and groupers, those animals are incapable of beliefs about the future. Therefore they are incapable of expecting future interactions of potential cooperation or defection. This means that whatever game strategy best promotes survival (whether tit-for-tat, or other more complex variants) does not depend on knowing that the game being played is an iterated one. The coding of the game strategy must be at the very low cognitive level of spontaneous, instinctive responses to particular situations; it cannot depend upon planning or foresight. Crocodiles are unable to reason, “Ziczacs and I have been engaged in a cooperative enterprise for awhile now, and I have every reason to believe that we will be business associates well into the future. Since our relationship has been so mutually advantageous so far, I choose not to eat a ziczac today.” When confronted with a case in which a ziczac bird is in its mouth, a crocodile must just lack the instinct to eat it.

Human beings are of course able to engage in higher-level reasoning about the likelihood of future interactions with other agents. These future-directed beliefs will form a feedback mechanism altering the strategy that one plays in prisoner’s dilemmas. For example, with my spouse or my colleagues at work I know full well that

<sup>17</sup> The *locus classicus* being (Axelrod 1984), esp. Chap. 5. See also Joyce (2006), Sect. 1.4 for a survey of more recent literature.

our interactions will be iterated, and I might well reason my long-term advantage will be best promoted if I am particularly cooperative (although not naively playing All C no matter what) and forgiving of the occasional defection. However, if natural selection has utilized game-theoretic reciprocity to build spontaneous intuitions or case-specific instincts in humans (as it has in nonhuman cooperators) then those intuitions are not going to be the result of sophisticated reasoning about the future. *Homo sapiens sapiens* has only been around for 200,000 years, and the entire *homo* line only for a couple million or so. This is not long enough for evolution to hardwire instincts that depend upon the advanced features of our new big brains.

There are good reasons to think that at least some of our instincts about altruism, deception and punishment, helping, reciprocity, and cooperation are the result of evolved game theoretic strategies, and I have given a gloss of some of this reasoning. The intuitions we should expect to be produced by these mechanisms are noninferential, spontaneous responses to particular cases.

Should we expect these intuitions to be agent-neutral, like those generated by the following case? *Sports Car*. A man is driving his new sports car when he sees a child at the side of the road with a bloody leg. The child asks the car driver to take her to the nearby hospital. If the driver helps the child, it will cost him \$200 to clean and restore the leather interior. In this case, most people judge that the driver ought to take the child to the hospital, despite the personal cost.<sup>18</sup> Kin selection isn't going to produce intuitions to help unfamiliar non-kin like the bleeding child in the Sports Car case. However, I believe that reciprocity likely will.

Recall that an agent-neutral moral theory is one according to which everyone has the same duties and moral aims, no matter what their personal interests or interpersonal relationships. Prisoner's Dilemma, and other interaction games, are abstract game theory puzzles. The solutions to them are mathematical, computable and are the solutions no matter who is playing the game. Computers in Prisoner's Dilemma tournaments never cohabitated with their opponents and don't even know who they are. Their strategies are programmed in and the machines play in response to the moves of the other players whomever they may be.

Now, one might object that we are not prepared to enter into prisoner's dilemma games (or other similar games such as chicken or stag hunt) with just any other creature, of course. Crocodiles may cooperate with ziczacs, but not just any bird that wanders into its maw. And it is a brave soul who is ready to be an initial cooperator and play a forgiving game of two-tits-for-a-tat with a cobra or a grizzly. It may be that after eons of playing asynchronous games that evolution has built in an aversion to cooperate with certain species,<sup>19</sup> or perhaps we have collectively learned through bitter experience that cobras are repeat defectors and have passed the knowledge down through the generations. Either way we learn to separate reciprocators/cooperators from nonreciprocators/defectors, and rudimentary notions of fairness, cheating, and

<sup>18</sup> I borrowed this example from Hauser (2006, p.61).

<sup>19</sup> The literature on asynchronous and evolutionary prisoner's dilemma strategies is interesting and complex. See the excellent review article Kuhn, Steven, "Prisoner's Dilemma", *The Stanford Encyclopedia of Philosophy* (Fall 2003 Edition), Edward N. Zalta (Ed.), URL = <<http://www.plato.stanford.edu/archives/fall2003/entries/prisoner-dilemma/>>.

free-riding can begin to arise. Even though reciprocity may eventually yield discrimination between groups of players that are recidivist defectors and those that are not, such partitioning is not coextensive with kin selection. Crocodiles and ziczacs don't need to have been raised in the same home in order to engage in reciprocal helping behaviors, they don't need to be kin or pass for kin. And I don't need to personally know the bleeding child on the side of the road to feel the moral pull of a request to aid her, even at some personal cost. While it is clear that there are other factors and forces at work in bounding the class of those to whom we feel some moral connection, game theory alone is wholly neutral about who plays the game. Insofar as game-theoretic reciprocity is responsible for moral intuitions, we should expect those intuitions to be agent neutral ones.

## 5 Relativism and skepticism

To this point I have argued that natural selection produces two different kinds of moral intuitions: agent-centered intuitions resulting from kin selection, and agent-neutral ones arising from reciprocity. James Dreier has argued that the agent-neutral/agent-centered dichotomy is the primary taxonomic division between classes of normative theories (Dreier 1993). Typically deontology and virtue theory align under agent-centered theories<sup>20</sup> and egoism and consequentialism are considered kinds of agent-neutral theories. Dreier, however, thinks any moral theory can be recast as a maximizing theory. Traditional hedonistic utilitarianism, which takes pleasure as the highest good, tells us to maximize the amount of pleasure in the world. On the face of it, deontology is quite dissimilar, since it says that we have moral rights that need to be respected regardless of the consequences of our actions (e.g., you shouldn't kill the innocent, even if it would produce more net pleasure in the world). Dreier argues that even deontology can be considered maximizing—it's just that instead of maximizing pleasure and minimizing pain, we should maximize the respecting of rights and minimize rights violations. Yet the fact that any given moral theory can be viewed as a maximizing one (and in that sense, a consequentialist one) doesn't make every theory agent-neutral. Put another way: classical utilitarianism is both maximizing and agent-neutral. But the maximizing feature is not what distinguishes it from deontology or from other moral theories. Agent neutrality is. Therefore it is agent-neutrality and agent-centeredness that is the root division among moral theories. Whether Dreier is right about this or not, it is clear that the agent-neutral/agent-centered split marks an important divide in our moral thinking.<sup>21</sup> Given the common moral epistemology outlined earlier, it is therefore unsurprising that our most prominent moral theories should bifurcate into agent-neutral and agent-centered ones.

Since the agent-neutral/agent-centered rift is rooted in our most basic moral intuitions and indeed in our very nature as creatures with moral instincts, there is no recourse to an epistemically more fundamental stance to resolve the dispute. Since

<sup>20</sup> Cf. Harris (1999).

<sup>21</sup> It is an important division even if, like Richard Brook, one thinks that many (although not all) deontological proscriptions have agent-neutral formulations. See Brook (1997).

agent centered and agent neutral theories seem equally coherent and internally consistent and we can get no deeper than our original data, which is our spontaneous reactions to cases, we have every reason to believe that the dispute between agent centered and agent neutral theories is irresolvable. As a result, we can either be skeptics (like Joyce)—morality is nothing more than a story we make up as a result of the evolutionary encoding of certain instincts and has nothing to do with truth or reality—or we can be relativists. There are moral truths, and we can have knowledge of these truths. However, moral truths (and concomitantly our knowledge of them) is relative to doxastic perspective; that is, it is relative to the data we use to develop our moral theory. Agent-neutral theories are true, relative to the perspective of the agent neutral intuitions about cases that are used as the foundational basis for those theories. Agent centered-theories are likewise true, *mutantis mutantis*.

I have no general argument against the skeptical option except to note this: the only chance for any sort of moral realism, with true ethical propositions and genuine moral properties, is relativism. As distasteful as relativism is for most ethicists, it is still more palatable than skepticism. The full details of moral relativism, the nature of ethical perspectives, the relations among different perspectives, and so on, cannot be worked out here.<sup>22</sup> Suffice to say, though, that there is considerable motivation to flesh out those details.

The usual objection to moral relativism is that it makes unjust all moral criticism of the ethical views in alternative perspectives. Here is a familiar version of the argument from Mary Midgely with respect to a moral relativism that relativizes the truth of moral propositions to cultures.<sup>23</sup> If moral relativism is true, then any moral statement is true in one culture and false in another. If a moral statement is true in one culture and false in another, then it is impossible for a person in the first culture to justly criticize the moral beliefs of a person in the second culture. For example, it would be impossible for us to criticize accurately the medieval Samurai for their practice of slicing wayfarers in half to test their new *katanas* or to condemn the Nazis for the Holocaust. Such a judgment would be akin to Mary judging that Bob's utterance of "I am a man" is false because it is false when Mary says it. Nor could we even legitimately critique the moral practices of our own society, since *ex hypothesi* the moral truth (relative to our society) is whatever is true in our society's perspective and so to criticize it from within the society is foolishly to criticize the truth. By the same reasoning, it would also be impossible for us accurately to praise Samurai culture, Nazi society, or ourselves. In short, we cannot coherently make judgments at all about the moral codes, standards, or behavior of people in any culture at all. But we do coherently judge both other cultures and our own; therefore, moral relativism is false.

As an argument against simple cultural ethical relativism, Midgely's argument is appealing. Behind her idea that we can evaluate the moral beliefs and theories in other cultures (and that they can likewise evaluate ours) is the assumption that there is neutral rational stance from which all interested parties can deliberate. That is, we can assemble a delegation from each of the world's cultures and subcultures, like an

<sup>22</sup> Although I have much to say about how to develop relativist theories in Hales (2006).

<sup>23</sup> Midgely (1981, Chap. 5).

United Nations of ethical theory, and reason together—critiquing each other, revising our views, negotiating, considering cases. And in the end we hammer out an agreement that all the negotiators hold in reflective equilibrium. However, if I am right about the plurality problem and there is a schism in our moral intuitions at the most basic level, ultimately resulting in equally coherent yet incompatible moral theories, then Midgely’s UN of morality cannot possibly succeed. There is no stance behind our most basic intuitions, no additional place for reason to go, and so the possibility of universal agreement in reflective equilibrium is at most vanishingly small.

In an example of the canard that one person’s ponens is another’s tollens, Peter Singer accepts my conclusion that moral intuitions are the consequence of our evolutionary history and that intuitions so produced are unlikely to be consistent ones (Singer 2005). We both accept this conditional: If moral theorizing proceeds by taking intuitions as *prima facie* data and then attempts to develop more comprehensive theories that can be held in reflective equilibrium and fundamental moral intuitions are ineluctably inconsistent, then we should be either moral skeptics or moral relativists. I suggest that relativism is a more attractive option than skepticism. Singer instead takes the argument as a *modus tollens*—skepticism and relativism must be wrong, but since the evolutionary psychological account of moral intuition is right, what must be rejected is the traditional model of moral epistemology.<sup>24</sup> However we come to have moral knowledge, it isn’t through intuition or through the development of theory by reflective equilibrium.

The chief problem with Singer’s *modus tollens* is that he fails to offer a cogent alternative model of metaethics or ethical epistemology. He gives an example of a person who overcomes an initial moral intuition that they should not push a stranger in front of a runaway trolley in order to stop it. In overcoming this initial judgment, the person reasons that (1) she would divert the trolley onto a side track and thus kill one person walking on the side track to save five aboard the trolley, (2) the first case is just like the second, (3) one ought to treat like cases alike, and (4) the death of one person is a lesser tragedy than the death of five. Singer thinks that the traditional model of intuitions taken as *prima facie* data and then systematized by reflective equilibrium cannot explain the preceding example. Singer’s view here is odd on the face of it—the abandoning of an initial intuition after due consideration and weighing against other intuitions and principles to which one is committed just *is* the method of reflective equilibrium.

Moreover, Singer acknowledges that the only reason to believe that one ought to treat like cases alike or that the death of one is a lesser tragedy than the death of five is intuition, but he strangely claims that this intuition “does not seem to be one that is the outcome of our evolutionary past” Singer (2005, p. 350). Rather, he makes a distinction between moral intuitions (explainable by evolutionary psychology) and rational intuitions (not explainable by EP). Naturally, Singer takes his own utilitarian, agent-neutral intuitions to fall into the latter category. Not only is it hard to see this as anything but a case of special pleading, but I have previously given an argument as to how natural selection would also provide us with agent-neutral intuitions. Thus I

<sup>24</sup> Singer (2005); see esp. pp. 346–349.

conclude that, absent a genuine alternative model of moral knowledge, we should stick with my original conclusion: our choice is either moral skepticism or moral relativism.

## 6 Conclusion: a pragmatic relativism

I have argued that there is no means of justifying one fundamental theory over another, as all such justifications are ultimately grounded in root intuitions that diverge at the base. The upshot is that either we should be skeptics about moral truths or relativists. Either there is no reason to think that morality has anything more to do with reality than ghosts and witches, or moral propositions can be true, just not uniquely true. If we reject skepticism, how shall we advocate non-arbitrarily for any particular moral theory? The answer cannot be a normative one; that is, arguments that attempt to show that following one moral theory is likelier to lead to right action, or that the maxims and principles of one theory are more probably true, are doomed to failure. So one cannot promote, say, agent-centered morality because it is *better* in any moral or alethic sense.

What remains is the *pragmatic* endorsement of a moral theory. The judgment of which moral theory we ought to prefer must also be made on the basis of pragmatic, not normative, reasons. To offer normative reasons for selecting a moral system is to land us right back in the problem of intractable moral debates bogged down in the construction of ever more abstract and bizarre cases and counterexamples.

In this respect too morality is like languages. The choice of a language must be for reasons outside of communication, or expressive power; the reasons for thinking that some natural languages are more semantically limited than others are not impressive. Nevertheless, some languages are superior to others along various pragmatic dimensions. In the case of written languages, Arabic numerals are superior to Roman ones for purposes of mathematics (quick, what's MMMMCCCLVI ÷ CXXIX?) and Roman letters are superior to Chinese ideograms for keyboard typing. With spoken language, Sanskrit is the best language for punning,<sup>25</sup> and Italian, with many words ending in vowels, is better for opera than are languages with frequent fricatives and glottal stops at the ends of words.

What kinds of reasons might a pragmatist advocate for the selection of a moral theory? It seems to me that at least some of the nonmoral values that arise in discussions of theory selection in the natural sciences could be a beginning point. Values such as simplicity, elegance, and comprehensiveness could be desiderata. Moral philosophy has one additional quality that natural science does not—it is supposed to be *practical*. If scientific theories have a practical side, then that is something of a bonus. But morality exists specifically to be action-guiding; should it lose that quality it becomes difficult to see why we should care about it. So another value to consider in the choice of a ultimate moral theory is its *usefulness*. A byzantine, cumbersome, impractical theory may be rejected in favor of one that, even if no better at accommodating our intuitions, can be easily and simply applied.

<sup>25</sup> On the advantages of Sanskrit for punditry, see Ostler (2005, p. 184).

Perfect human reasoners in perfect reflective equilibrium will never settle upon one moral theory using moral intuitions as data. As we have seen, the very generation of those intuitions is bound up with the evolutionary history of our species and even, taking a wider view, the evolutionary history of life on Earth. Furthermore, our intuitions diverge into agent-centered and agent-neutral right at the very start, making the ancient rationalist hope of rapprochement a pipe dream. If we think that moral propositions have truth values at all—that is, if we reject skepticism—then those truth values must be relative to which doxastic perspective we adopt, one grounded in agent-centered intuitions, or one grounded in agent-neutral ones. Even further, if we are to promote one particular moral theory, it cannot be because that theory has a better claim to the truth. Rather, it must be because that theory is in some way more practical for our needs. Faced with a plurality of moral theories equally consonant in reflective equilibrium—all equally true by relativist lights—we should pick among them on the basis of non-moral pragmatic virtues.

**Acknowledgements** The author presented a distant ancestor of this paper at The Philosophy of Pragmatism conference at Babes-Bolyai University in Cluj, Romania. Thanks to Gary Hardcastle for comments on the penultimate version. Tim Johnson and Richard Brook provided helpful and relentless criticism from start to finish, and I appreciate the tough love.

## References

- Axelrod, R. (1984). *The evolution of cooperation*. New York: Basic Books.
- Brook, R. (1997). Is Smith obligated that (she) not kill the innocent or that she (not kill the innocent): Expressions and rationales for deontological constraints. *The Southern Journal of Philosophy*, 35(4), 451–461.
- Brown, D. E. (1991). *Human universals*. New York: McGraw-Hill Publishing Co.
- Darwin, C. (1882). *The descent of man and selection in relation to sex* (2nd ed.). London: John Murray.
- Dawkins, R. (2006). *The god delusion*. New York: Houghton Mifflin.
- de Waal, F. (2006). *Primates and philosophers: How morality evolved*. Princeton: Princeton University Press.
- Dreier, J. (1993). Structures of normative theories. *The Monist*, 76, 22–40.
- Flack, J. C., & de Waal, F. B. M. (2000). 'Any animal whatever': Darwinian building blocks of morality in monkeys and apes. In L. D. Katz (Ed.), *Evolutionary origins of morality: Cross-disciplinary perspectives*. Bowling Green, OH: Imprint Academic.
- Gazzaniga, M. S. 2005. *The ethical brain*. New York: Dana Press.
- Hales, S. D. (2006). *Relativism and the foundations of philosophy*. Cambridge, MA: MIT Press (A Bradford Book).
- Harris, G. W. (1999). *Agent-centered morality*. Berkeley: University of California Press.
- Hauser, M. D. (2006). *Moral minds: How nature designed our universal sense of right and wrong*. New York: HarperCollins.
- Hintikka, J. (1999). The emperor's new intuitions. *The Journal of Philosophy*, 96(3), 127–147.
- Hume, D. (1978 [1739]). In L. A. Selby-Bigge (Ed.), *A treatise of human nature* (2nd ed.). Oxford: Oxford University Press (Clarendon Press).
- Joyce, R. (2006). *The evolution of morality*. Cambridge: MIT Press (A Bradford Book).
- Lieberman, D. (2007). Moral sentiments relating to incest: Discerning adaptations from by-products. In W. Sinnott-Armstrong (Ed.), *Moral psychology, Vol. 1: The evolution of morality*. Cambridge: MIT Press.
- Lieberman, D., Tooby, J., & Cosmides, L. (2003). Does morality have a biological basis? An empirical test of the factors governing moral sentiments relating to incest. *Proceedings of the Royal Society: Biological Sciences*, 270, 819–826.
- Lowry, R. (2006). The F world. *National Review Online*, August 8.



- Mackie, J. L. (1977). *Ethics: Inventing right and wrong*. New York: Penguin Books.
- McMahan, J. (2000). Moral intuition. In H. LaFollette (Ed.), *The Blackwell guide to ethical theory*. Oxford: Blackwell.
- Midgely, M. (1981). *Heart and mind: The varieties of moral experience*. New York: St. Martin's Press.
- Ostler, N. (2005). *Empires of the word: A language history of the world*. New York: HarperCollins.
- Pinker, S. (1994). *The language instinct*. New York: William Morrow and Company.
- Ridley, M. (1996). *The origins of virtue: Human instincts and the evolution of cooperation*. New York: Viking Penguin.
- Russell, B. (1927). Why I am not a Christian. In R. E. Egner & L. E. Denonn (Eds.), *The basic writings of Bertrand Russell*. New York: Simon and Schuster.
- Shermer, M. (2004). *The science of good and evil*. New York: Henry Holt.
- Singer, P. (2005). Ethics and intuitions. *The Journal of Ethics*, 9, 331–352.
- Sober, E., & Wilson, D. S. (1998). *Unto others: The evolution and psychology of unselfish behavior*. Cambridge: Harvard University Press.
- Trivers, R. L. (1971). The evolution of reciprocal altruism. *The Quarterly Review of Biology*, 46(1), 35–57.
- Wilson, E. O. (1998). *Consilience*. New York: Alfred A. Knopf.
- Wright, R. (1994). *The moral animal*. New York: Pantheon Books.