

***Moral Instincts and the Problem with Reductionism:***

***A critical look at the work of Marc Hauser***

**Joseph Bankard**

**Assistant Professor of Philosophy**

**Northwest Nazarene University**

In his recent book *Moral Minds*<sup>1</sup>, Marc Hauser argues that many of our moral instincts have a biological foundation. In the past, such a claim would have received heavy criticism from both religious and scientific circles because morality, understood as a biological adaptation, seems to lose much of its normative weight. However, current work in sociobiology and evolutionary psychology strongly supports a biologically based morality that need not do serious damage to morality's ontological status.

In light of the exciting research being done in evolutionary ethics, this paper seeks to accomplish two things. First, I intend to use the work of Marc Hauser to argue that human morality does have a clear biological foundation. The data is simply too strong to ignore.

Second, I intend to argue that while Hauser provides a strong case for a biologically based morality, some of his conclusions participate in a form of "greedy reductionism." The problem, as I see it, is Hauser acknowledges the bottom-up influence taking place between biology and moral judgments, but he refuses to account for the top-down influences that occur between metaphysical assumptions, moral theories, and human behavior. The current critique will focus on Hauser's failure to provide an adequate framework for justifying our moral judgments. It is one thing to argue that some of our fundamental moral instincts are biologically inherited, but it is quite another to justify moral judgments at the instinctual level. Hauser's attempt to reduce moral judgments to the level of instinct fails to provide a sufficient context for moral justification revealing a form of greedy reductionism.

In *Moral Minds*, Hauser suggests that many of our moral instincts are inherited through our evolutionary past. Thus, the foundation of human morality is our shared

biology. To help make this argument, Hauser highlights the similarities between language acquisition and the acquisition of morality. Hauser argues that while humans learn a particular language through training and environment, the biological components necessary for language acquisition are universal. Simply put, humans are biologically designed to develop language.

According to Noam Chomsky and others, there exists a wide variation with respect to language, but all languages are grounded in basic principles of grammar that are universal. For example, all languages use certain *word order* parameters such as the word placement used to form a question as opposed to the word order used to form a proposition. The word order for propositions and questions varies among languages; however, no language uses the same word order for each. This is a universal constraint placed on all languages without exception. Thus, there is a universal grammatical constraint on the variation found in languages.

Similarly, in all languages there is a distinction made between words classified as nouns and words classified as verbs. For example, most languages use either a subject-verb-object or a subject-object-verb word order. If there were no linguistic parameters limiting the range of variation, then we should expect to find languages with all kinds of word orders, but this is not the case. It appears that there is an unconscious logical grammar that causes humans to generate languages that follow certain universal constraints.

Hauser writes,

If the principles-and-parameters approach is right, at least in some form, then we can characterize linguistic diversity by identifying how each of the core parameters is set. Flip a few parameters this way and that, and you get English. Flip the same ones a different

way, add a few more parameters, and you get Chinese. If this approach is correct, then we can not only explain the diversity of all possible languages, we can also identify a set of impossible languages, those that are not learnable based on the constraints set up by the innate principles that constitute our universal grammar.<sup>2</sup>

Moral development parallels language in this way. Thus, one may inherit a particular moral framework from one's environment, but the biological components necessary for moral development are universal. Again, it looks as though humans are biologically designed to acquire morality.

Hauser writes,

Why does everyone take for granted that we don't learn to grow arms, but rather, are designed to grow arms? Similarly, we should conclude that in the case of the development of moral systems, there's a biological endowment which in effect requires us to develop a system of moral judgment and a theory of justice, if you like, that in fact has detailed applicability over an enormous range.<sup>3</sup>

The thrust of Hauser's argument is that an unconscious set of moral instincts has evolved in humans creating a universal moral grammar. This moral grammar is not rigid in nature, but rather, provides a set of general moral parameters that allow for a significant amount of flexibility and variation. It is important to note that Hauser's use of innate morality differs greatly from a thinker like Plato. Unlike Plato and other rationalists, Hauser's moral grammar does not provide the individual with specific moral content. Instead, Hauser argues that humans are born with abstract rules and principles that function as parameters on our moral systems. This view of innate morality suggests that humans are born with the necessary equipment to develop moral systems, but the specific moral content that develops is the result of one's environment. Thus, one's innate moral grammar provides the universal parameters and principles for our moral systems, but these parameters allow for a wide variation with regard to content.<sup>4</sup>

...Like language, the specifically expressed and culturally variable moral systems are learned in the sense that the detailed contents of particular social norms are acquired by exposure to the local culture; the abstract principles and parameters are innate. The role of experience is to instruct the innate system, pruning the range of possible moral systems down to one distinctive moral signature.<sup>5</sup>

In this way, Hauser resembles E.O. Wilson because both thinkers argue that biology holds culture on a sort of leash, giving it some room to roam, while at the same time providing strong limits that cannot be traversed. However, unlike Wilson, Hauser gives much more formative power to one's culture and one's environment. A person's moral awareness is strongly influenced by culture, religion, upbringing, and the like, but it would be a mistake to say that morality is nothing more than one's environment. Underneath all the cultural variation is an unconscious moral grammar that is both universal and innate.

For example, not all people will come to the exact same conclusions regarding intentional versus unintentional killing. But all people will make a distinction between intentional and unintentional killing. Not only will this distinction be made, but it will be seen as morally relevant as well. The universal nature of such judgments leads Hauser to conclude that they must stem from a biologically based moral grammar.

In order to support his overall argument, Hauser must do three things. First, he must provide solid evidence supporting the existence of a universal moral grammar. Second, Hauser needs to show the relevant connections between this moral grammar, moral judgments, and human behavior. Finally, Hauser must provide a framework whereby our moral grammar and subsequent judgments can be justified. If Hauser can defend each of these points, then a strong case is made supporting a biologically-based moral grammar.

### *Moral Instincts*

In order to support the existence of an innate moral grammar, Hauser makes a shift from moral emotions and moral reasoning to moral instincts. To highlight this transition, Hauser describes three different moral creatures: the Humean creature, the Kantian creature, and the Rawlsian creature. The Humean creature focuses on moral emotions as opposed to moral reasoning. Hume states that “Reason is and ought only to be the slave of the passions, and can never pretend to any other office than to serve and obey them.”<sup>6</sup> For Hume, moral decision making is not the result of rational thought. Rather, moral decisions are made based on emotion. Hume argued that those actions that produce good feelings such as pleasure, pride, satisfaction, and the like, were said to be “good,” while those actions that produced pain, guilt, and fear were said to be “bad.” These conclusions are not the product of reason. Rather, reason is used after the fact as a way of justifying our moral emotions.

Kant, on the other hand, argued that emotions were too subjective to be used in our moral evaluations. If morality is to be just and fair, emotions would have to be put aside in order for reason to bring us to universal conclusions. Instead of using emotions to make moral decisions, one should evaluate a decision rationally to see if a given action can be universalized. The process of universalizing moral norms is used as a way of uncovering one’s biases and selfish motives. Once these are revealed, a moral decision can be made that is rational and just.<sup>7</sup>

Hauser argues that while both the Humean and Kantian creatures are important for moral judgments and moral actions, both are fundamentally flawed as well. Moral judgments do not originate with one’s emotions nor do they result from rational

deliberation. Rather, our most fundamental moral judgments are the result of unconscious moral instincts. Hauser argues that most moral judgments happen immediately before reason or emotions are raised to the level of conscious thought. Many times the moral judgments that are made happen without the agent actually being aware of any specific moral principle or moral emotion at work. This leads Hauser to conclude that such judgments are the result of unconscious moral instincts. After the initial moral judgment is made, both reason and emotion play a large role in systematizing moral thoughts and actions. However, this only happens after an unconscious moral evaluation has been made.

It is true that emotions strongly influence human behavior, but emotions must be triggered by something prior. Hauser argues that moral instincts lead to unconscious moral judgments. These moral judgments trigger emotions. Similarly, reason begins its evaluative process only after the situation has been unconsciously appraised by one's moral grammar. Thus, both reason and emotion arise as a result of a prior moral evaluation. Hauser refers to this type of moral agent as the Rawlsian creature.

Hauser writes,

Rawlsian creatures are appraisers, but their appraisals are unconscious and emotionally cool. The appraisal can either trigger a feeling or not. The appraisal triggers an emotion before a judgment is delivered. Alternatively, the appraisal triggers the judgment, which then triggers an emotion.<sup>8</sup>

Much of the evidence used by Hauser to support the Rawlsian creature was gathered from his "Moral Sense Test." The test includes moral dilemmas such as trolley examples and game theory examples that have been taken by more than sixty thousand participants on-line.<sup>9</sup> The sample is diverse in age, gender, religious conviction, socio-

economic status, and the like. Across all of these demographics, the results are very consistent. Moreover, the Moral Sense Test has now been translated into five different languages including Hebrew, Arabic, Indonesian, Chinese, and Spanish. Once again, the results were very consistent across a broad demographic.<sup>10</sup>

One of the moral dilemmas has a man driving a new car with leather interior. There is a child on the side of the road with a bloody leg calling out for help. The driver realizes that the child needs immediate medical assistance, but her bloody leg will most likely ruin the new leather seats resulting in a 200 dollar expense. What should the driver do? After reading the scenario, participants were asked if helping the child was morally permissible, obligatory, or forbidden. Does the driver have a moral obligation to help the child or is it merely morally permissible? What is your immediate response to a person who drives past a hurting child in order to protect his leather seats?

The second scenario has the same man sorting through his mail after work. On this particular day UNICEF has sent a letter asking for a 50 dollar donation that will result in twenty five children receiving oral rehydration salts that could potentially save their lives. Should the man give 50 dollars to UNICEF? Does the man have an obligation to send the money or is it merely permissible? What is your immediate response to a person who reads the letter and throws it away without giving a donation?

Hauser points out that when given these two examples, participants almost immediately and without hesitation answer that the man has a moral obligation to help the child with the hurt leg, but there is no obligation to send the money to UNICEF. Similarly, participants have a stronger emotional reaction to the man who drives by the injured child in an attempt to save his leather seats, than the man who throws the

UNICEF letter away. Why are such reactions so common? It isn't the result of rational deliberation because looking at the two examples rationally will show that the UNICEF example requires less money, but the end result will be twenty five lives saved instead of just one. Shouldn't there be a stronger moral obligation to spend the 50 dollars to save the twenty five African children, rather than, spending 200 dollars to help just one injured child? Certainly the lives of African children are just as valuable as the life of the girl on the side of the road, so shouldn't the same obligation hold?

Hauser uses this example to show that people respond to examples such as these out of instinct and not due to some rational deliberation. The responses are often immediate and almost no participant can give a strong reason for their answers. Thus, the moral evaluation that takes place seems to be unconscious and instantaneous, leading Hauser to conclude that it is part of humanity's universal moral grammar. Hauser suggests that there does seem to be a moral distinction being made, but this distinction never reaches the level of consciousness. Rather, it operates more like a moral instinct than a moral principle. According to Hauser, the distinction being made is between moral actions and moral omissions. The man driving the car has to take direct action in order to avoid the suffering child, whereas the man who throws away the UNICEF letter simply avoids sending money. One represents a moral action, while the other represents a moral omission. Hauser argues that the action/omission distinction is a universal moral instinct that reveals itself in numerous examples. Across the board, people tend to judge moral actions more harshly than moral omissions. For example, we are more likely to judge an action with negative consequences as forbidden whereas we judge the omission of an action with the same negative consequences as permissible.<sup>11</sup>

Instead of turning to reason to explain such moral instincts, Hauser argues that we should turn to our evolutionary history. Since many of our moral instincts would have developed thousands of years ago, we must look at this early environment to help illuminate our present condition. In the past, it would have been impossible to help people at a distance. All possible forms of altruism were those that occurred within one's immediate sphere of influence. Thus, when we encounter immediate suffering we perceive a stronger moral obligation than when we hear about the suffering of others at a distance. If our moral instincts developed during a time when long distance altruism was not possible, then this would help explain such an instinct. This leads Hauser to articulate, what he considers to be, a universal moral instinct.

If we can *directly* prevent, with a high degree of certainty, something bad without sacrificing anything of comparable significance, we are obligated to do it.<sup>12</sup>

Another evolutionary influence might be the distinction between in-group versus out-group altruism. It is very common to see in-group altruism in non-human animals, but it is rare to find examples of out-group altruism. If our moral instincts are the product of our evolutionary past, then humans should also display greater levels of altruism for in-group members than we do for out-group members. In the two examples above, there is a much greater chance that the girl hurting on the side of the road is an in-group member than are the individuals being helped by UNICEF. This in-group/out-group distinction might also contribute to the responses given.

The responses to the UNICEF dilemma support Hauser's claim that moral evaluations occur before rational deliberation. This poses serious problems for the Kantian creature. However, one could argue that the Humean creature does provide a

sufficient explanation for the data. Driving past a bleeding child creates a more visceral reaction than does the UNICEF letter. In one scene there is personal contact and a screaming child. In the other, there is a letter describing the need, but this cannot generate the same kind of urgency. For this reason, one might argue that the responses given to the UNICEF dilemma are the result of moral emotions, not moral instincts. It seems quite possible that certain moral actions create a stronger emotional reaction in most people than do moral omissions, thus explaining why it seems worse to drive by a suffering child than to throw away a UNICEF letter.

In order to respond to such objections, Hauser raises a second set of examples taken from Jonathan Haidt.<sup>13</sup> In one scenario, a man comes home to find his daughter in tears. “What’s wrong sweetheart?” the father asks. The daughter proceeds to tell him that she came home from school to find her mother eating ice cream off of the toilet seat. Participants are then asked to rank this example on a disgusting scale from 1 to 10. Typically this example scores pretty high (average is about a 7).

The second scenario has a brother and sister on vacation. The two have a wonderful relationship and they decide to consummate it by making love. The brother has had a vasectomy and the sister is on birth control, so there is no chance of pregnancy. The two make passionate love and keep it a lifelong secret as something that they both will remember and cherish. Again, participants are asked to rank this on a scale from 1 to 10. As you might expect, this example scores very high on the scale (typically an 8 or 9).

Both of these examples illicit a similar emotional response, disgust! However, Hauser argues that there is a fundamental distinction that almost all people participating in the survey make. The first example is considered disgusting, but few people argued

that it was morally wrong. It seems that there are situations that might help justify the mother's behavior such as a freshly washed toilet seat or perhaps there is some cultural difference that might help explain the behavior.

In the second scenario, there doesn't seem to be any potential justification for this kind of incest. Not only does such an example seem disgusting, but it seems morally deficient as well. Why do almost all people who participate in these surveys make this same distinction? Hauser argues that this is yet another aspect of our universal moral grammar.

If Hume is right and emotions dominate our moral judgments, then the disgust experienced in the two examples should lead to similar moral conclusions. As we have seen, this is not the case. It seems that something else is going on that enables us to make a distinction between those things that are merely disgusting and those things that are both disgusting and immoral. Despite the similarity in our emotional responses to these two scenarios, our moral instincts make a clear distinction between the two. According to Hauser, this distinction is the result of our unconscious moral grammar.

**(More on Moral Instincts?)**

### *Connecting Moral Instincts to Moral Judgments*

So far, Hauser has provided a strong defense of moral instincts in the form of the Rawlsian creature. The examples and analysis given make a strong case for a moral grammar that helps make moral evaluations and judgments before either reason or emotion play a strong role in decision making. The next step is for Hauser to show the connection between our unconscious moral instincts and our conscious moral judgments.

One way Hauser reveals this connection is through trolley examples made famous by Philippa Foot, which highlight the distinction between killing and letting die. The dilemma uses two different scenarios involving Denise, Frank, and a Red Trolley. In each situation, Denise or Frank must make a decision that will result in the death of one person or the death of five people. After each scenario, participants were once again asked whether the actions performed were morally permissible, obligatory, or forbidden. It should come as no surprise, that the results were quite consistent across a broad demographic.

In the first scene, Denise is standing near a railroad track when she notices a trolley that is out of control. If the trolley continues on its course, it will crush five hikers on the tracks in front of it. Denise has an opportunity to pull a level that will send the trolley down an alternate track where one hiker is standing and will surely be killed. What should Denise do? Is Denise morally obligated, merely permitted, or absolutely forbidden from pulling the lever?

In the second scene Frank is standing on an overpass watching the trolley race towards the five hikers on the track. Frank realizes that the five hikers will be killed unless a large weight is thrown in front of the trolley in an attempt to stop it. Luckily, there is a heavy set man standing next to Frank on the over pass. What should Frank do? Is Frank morally obligated, permitted, or forbidden from throwing the man onto the tracks?<sup>14</sup>

The overwhelming answer given to the first scenario is that Denise is morally obligated to pull the lever in order to save five hikers. However, there was a similar consensus suggesting that at best Frank is morally permitted to throw the man onto the

tracks, although most participants answered that pushing the man was forbidden. Why is there such a unified discrepancy between the moral evaluations of these two dilemmas? When asked for the reasons behind their answers, many of the people surveyed gave a utilitarian response for the first dilemma involving Denise. However, this explanation does not seem to hold for Frank's dilemma. Here Frank is going to push one man off of the platform in order to save the lives of five others which should satisfy the principle of utility, but those who gave a utilitarian response to Denise's action were unwilling to provide the same justification for Frank. Similarly, those who tried to give a deontological justification for judging Frank's action to be immoral were unwilling to apply the same deontological principle against killing to Denise's situation. Thus, it seems that people are confident in their moral judgments, but seem to be relatively unaware of the reasons behind these judgments.

Hauser writes,

Thus, the intuitive knowledge underlying our moral judgments is like the intuitive knowledge of language, physics, psychology, biology, and music... We know that two solid objects can't occupy the same space at the same time and that a solid object will fall straight down if unsupported and unobstructed on the way down. But we know these factoids in the absence of a course in physics and without being aware of them. Mikhail's claim, and the key idea driving my argument for the moral faculty, is that much of our knowledge of morality is similarly intuitive, based on unconscious and inaccessible principles for guiding judgments of permissibility.<sup>15</sup>

Turning back to our examples of Denise and Frank, what moral instincts, if any, lead to the near universal judgments made regarding their dilemma? It is important to note that the two examples are similar in that both Denise and Frank must make a decision that will result in one death or five deaths, but there are some significant differences as well. For one thing, Denise's act is impersonal and neutral (flipping a

switch), whereas Frank's act is personal and negative (pushing a man off of a platform). Also, the negative consequences of Denise flipping the switch (one hiker dying) are merely foreseen, while the consequences of Frank's action (the large man falling to his death) are intended. These differences are not insignificant and might actually provide some insight into the unconscious moral judgments being made.

Hauser focuses on the difference between intentional killing and killing as the foreseeable and undesirable consequence of an action intended to produce a greater good. It seems as though people surveyed responded negatively to Frank because he must intentionally push a man off the overpass in order to stop the train. This action is immediately judged to be morally prohibited because it is an intentional act of harm, whereas, Denise is permitted to pull the switch because the goal of the action is to save five hikers not to kill the one standing on the side track. The hiker on the side dies as a foreseeable consequence of the action, but it is not intended. This is referred to as the principle of double effect.

Double effect is a traditional moral and legal principle... according to which otherwise prohibited acts may be justified if the harm they cause is not intentional and the act's foreseeable and intended good effects outweigh its foreseeable bad effects.<sup>16</sup>

The responses to the trolley examples are nearly universal, suggesting that the moral evaluations given are unconscious and innate; unconscious because there is no clear reason for the difference other than one action "seems" right while the other "seems" wrong and innate because this moral judgment appears to be universal. Hauser argues that the underlying moral principle guiding these moral judgments is if harm is the by-product of achieving the greater good, then it is morally permissible. However, if harm is the means to achieving the greater good, then it is morally forbidden.

The results of Hauser's moral sense test confirm that an unconscious distinction is often made between intentional harm and foreseeable harm. This strongly suggests that some of our innate moral principles have a biological foundation. This is not to suggest that answers to dilemmas like the trolley example will be completely universal. There will always be some variation with respect to culture, upbringing, and environment. However, underneath this variation is a set of moral instincts that limits the amount of variation, and leads to an almost universal distinction between intentional as opposed to unintentional killing.<sup>17</sup>

### *Moral Justification*

So far, Hauser has made a solid argument supporting the existence of an unconscious moral grammar that influences moral judgments and behavior. However, in spite of the strength of Hauser's research, certain problems still persist. For instance, in the trolley dilemmas Hauser attempts to draw two significant connections between biology and morality. The first is an attempt to connect the instinctual judgment that distinguishes intentional from unintentional behavior to the responses given for the trolley dilemma. Hauser argues that most participants side with Denise because the harm that she causes is unintentional. Frank, on the other hand, intentionally causes injury leading most participants to condemn his behavior. The connection that Hauser makes is strong because moral instincts should have a direct affect on moral judgments. If the distinction between intentional verses unintentional behavior is innate, then it will have a direct affect on one's immediate analysis of a situation like the trolley dilemma.

The second connection that Hauser makes is more dubious. After linking one's unconscious moral grammar to one's moral judgments, Hauser attempts to connect one's

moral judgments with the underlying moral principles used in such judgments as a means of justification. For instance, the initial judgment participants make regarding Denise and Frank results from the moral principle of double effect. Hauser argues that participants justify Denise's action because the harm caused is foreseeable, but not intended, while the harm caused by Frank is not only foreseeable, but also intended. Such a justification is common in moral and legal arguments.

A major problem with this aspect of Hauser's argument is almost none of the participants who provided responses to the trolley dilemma gave the principle of double effect as justification. This weakens Hauser's position because it points to a disconnect between the unconscious moral judgments that we make and the moral justifications that we give. It is one thing to claim that our initial moral judgments are the result of unconscious processes, it is quite another to argue that moral justification occurs as a result of these same processes. The reason most participants in the red trolley example were not able to identify the principle of double-effect as their justification for condoning Denise's behavior and condemning Frank's, is that each participant was operating at the level of instinct. However, moral justification does not occur at the instinctual level. For this reason, I argue that the principle of double effect cannot be justified simply because it is related to the innate distinction between intentional verses unintentional behavior. Rather, the principle of double effect can only be justified after it has withstood the test of genuine moral discourse.

Hauser has provided stellar evidence for universal moral instincts as well as valuable information related to their origin and application. Unfortunately, Hauser fails to provide an adequate framework for moral justification. It is one thing to claim that

certain innate principles affect moral evaluations as a result of their influence over moral development, but it is quite another to justify such principles based solely on their instinctual or biological status. It seems that before the evaluations that result from our innate moral grammar can be justified, a second level of moral discourse is needed. Discourse at this level must consciously evaluate our biological instincts and innate moral principles in an attempt to determine which ones should be followed and which ones should be ignored.

For example, favoring in-group members over out-group members with respect to one's affection and resources is a universal biological instinct.<sup>18</sup> However, the existence of such an instinct, by itself, is not enough to justify all forms of in-group favoritism. Before such behavior can be morally justified, the instincts and motivations influencing one's behavior must be analyzed and discussed to determine whether they fit a given moral standard. If certain biological instincts are deemed immoral, then it is our responsibility to act contrary to our biological nature. Such behavior is possible given the numerous examples of individuals who act in radically altruistic ways to intentionally benefit members of various out-groups.<sup>19</sup>

Clearly, there are two levels of discourse taking place in the example above. The first is a purely descriptive analysis of biological instincts and their influence over behavior. However, not all instincts that affect human behavior should be acted upon, making a second level of moral discourse necessary. Often after such deliberation, one concludes that while natural inclination favors one's in-group, morality demands that kindness and generosity be shown to one's out-group as well. Moral evaluations such as these are not usually the result of biological instinct, but rather, are the result of rational

discourse, deliberation, and intentionality. This ability to distance ourselves from our biological instincts in order to evaluate them reveals the transcendent nature of second level moral discourse.

Hauser has shown that in some very important respects biology gives birth to morality. All future moral discourse should be in serious dialogue with sociobiology and evolutionary psychology. However, once morality arises, it is able to transcend this biological foundation through second level moral discourse which seeks to evaluate the very innate principles that gave rise to it. By focusing solely on the bottom-up influence that biology has on moral judgments, Hauser has neglected the top-down influence that moral discourse can have on moral evaluations and actions. This does not happen by accident. I argue that Hauser neglects top-down causation because of certain core metaphysical convictions associated with naturalism, such as materialism and atheism. These metaphysical beliefs have a huge impact on Hauser's moral theory, but he never addresses them at the metaphysical level. Instead, these metaphysical conclusions are simply assumed. This represents a real limitation in Hauser's moral research program. By failing to provide a defense for metaphysical naturalism, Hauser has limited the scope of his research unnecessarily. The limited nature of Hauser's moral theory creates several problems, such as his refusal to consider the importance of top-down causation with regard to moral justification.

This point can be further highlighted by looking back at the example of in-group verses out-group membership. Hauser cites research done by Lawrence Hirschfeld<sup>20</sup> showing that children begin to make such distinctions at a very early age. One way that children do this is related to race. While Hauser acknowledges that race is a social

construction not a biological category, he also argues that such categories are part of our evolved psychology that helps identify in-group from out-group members. Hirschfeld showed that if young children are given pictures of two potential fathers, one white and one black, and then given a picture of a black baby dressed in clothing similar to the white father, the child will still be able to accurately identify the black man as the father. It seems that children are able to determine that skin color is more relevant than clothing in determining the baby's father. Clearly, young children do not have a sophisticated concept of race with heavy moral implications, but what this research does show is that young children are able to form innate categories associated with race.

While it is not part of a young child's nature to treat other racial groups poorly, over time, these innate categories of in-group and out-group affiliation often lead to unconscious racist behavior. Research done by the social psychologist Mahzarin Banaji<sup>21</sup> has revealed that adult individuals who claim to be inclusive and open-minded often make instinctual judgments that appear racist. For example, when participants are shown a rapid succession of faces and asked to judge each as good or bad, the results show a strong bias against all out-groups. Similarly, research regarding hiring practices reveals prejudices against hiring out-group members, even those with qualifications equivalent to competing in-group members. When confronted with the results of such research, the participants are usually shocked. It seems that such judgments are made unconsciously.

This research seems to support the existence of an unconscious moral grammar that influences our moral judgments and behavior. However, in this situation Hauser is not prepared to call the racist behavior that results from our innate grammar moral.

Hauser writes,

we form unconscious attitudes toward those in the out-group, and such prejudices lead to caricatures, this process is neither specific to the moral faculty nor operative in the same way... With respect to the underlying process, although our implicit attitudes operate outside the grasp of awareness, the principles are not at all inaccessible. They can be grasped by anyone looking carefully at their own patterns of judgments. And once we are aware of such biases, it is even possible to *override* them.<sup>22</sup>

It seems that Hauser is not prepared to generate a moral principle based on innate categories that distinguishes between in-group and out-group membership the way he did in the trolley example. However, it is difficult to see the difference between the unconscious distinctions that one makes concerning intentional versus unintentional killing and those related to in-group and out-group membership. Both are part of our unconscious moral grammar, both impact moral judgments, and both affect human behavior. In the trolley example Hauser justifies the distinction between intentional and unintentional killing using the moral principle of double-effect. Curiously, he is not willing to give the same kind of justification for the unconscious favoritism shown to in-group members calling it something to be “overridden.”

This strongly suggests that moral justification requires more than our unconscious moral grammar can provide. It seems that some innate categories lead to moral judgments, while others lead to immoral judgments. It is my contention that such distinctions require a second level of moral discourse. Hauser’s words also lend support to those who want to defend top-down causation. If moral discourse can help a person become aware of their in-group favoritism and enable them to overcome this prejudice, it reveals higher levels influencing lower levels. Unfortunately, Hauser fails to account for such causation in *Moral Minds*. At no point is it thoughtfully addressed. Moreover,

such causation seems difficult to explain given Hauser's explicit affirmation of metaphysical naturalism. It is my contention that one of the primary reasons for Hauser's refusal to seriously consider second-level moral discourse or top-down causation is due to these unnecessary metaphysical commitments.

The criticisms I have levied against Hauser do not diminish the major accomplishments of his research in *Moral Minds*. The evidence given by Hauser strongly suggests that there are unconscious principles that are both universal and moral. This research makes great strides toward an adequate form of universal naturalism due to the strong links between biology, evolution, and morality. What Hauser has failed to do is provide an argument showing how our moral judgments and behavior can be justified at the level of biological instinct. It is one thing to argue that certain innate principles play a role in our moral evaluations, but it is quite another to justify these principles based on their biological nature. My argument has been that one's unconscious instinct to judge a given situation cannot be considered moral or immoral until a second level of discourse is used to justify one's natural response.

---

<sup>1</sup> Marc Hauser, *Moral Minds: How Nature Designed our Universal Sense of Right and Wrong* (New York: Harper Collins Pub, 2006)

<sup>2</sup> Ibid, 74.

<sup>3</sup> Quote by Noam Chomsky in Marc Hauser, *Moral Minds*, preface.

<sup>4</sup> Ibid, 165.

<sup>5</sup> Ibid, 422.

<sup>6</sup> David Hume, *A Treatise on Human Nature* (Oxford: Oxford University Press, 1739), 474.

<sup>7</sup> For more on the Humean and Kantian creatures see Marc Hauser, *Moral Minds*, 22-38.

<sup>8</sup> Ibid, 192.

<sup>9</sup> The moral sense test can be found at [www.moral.wjh.harvard.edu](http://www.moral.wjh.harvard.edu).

- 
- <sup>10</sup> Hauser, *Moral Minds*, 127-129
- <sup>11</sup> See further research in *Moral Minds* pp. 8-20.
- <sup>12</sup> Ibid, 64.
- <sup>13</sup> Jonathan Haidt, "The Emotional Dog and its Rational Tail: A social intuitionist approach to moral judgment", *Psychological Review* 108(2001), 814-834.
- <sup>14</sup> Marc Hauser, *Moral Minds*, 114-116.
- <sup>15</sup> Ibid, 125.
- <sup>16</sup> Ibid, 124.
- <sup>17</sup> There are other examples cited by Hauser. Most notably, universal judgments regarding fairness and distribution.
- <sup>18</sup> See Richard Dawkins, *The Selfish Gene* (Oxford: Oxford University Press, 1976), W.D. Hamilton, "The Genetical Evolution of Social Behaviour, I and II", *Journal of Theoretical Biology* 7(1964), J.L. Hoogland, "Nepotism and Alarm Calls in the Black-Tailed Prairie Dog", *Animal Behavior*, 31(1983), G. Wilkinson, "Reciprocal Food Sharing in Vampire Bats", *Nature* 308(1984)
- <sup>19</sup> See *The Altruism Reader: selections from writings on Love, Religion, and Science*, ed. Thomas Jay Oord (Templeton Foundation Press: Pennsylvania, 2008), Lee Alan Dugatkin, *Cooperation Among Animals: An Evolutionary Perspective* (Oxford Press: New York, 1997)
- <sup>20</sup> Race in Mind (Hirschfeld, 1996); <http://www.umich.edu/~newsinfo/MT/96/Jun96/mta1j96.html>.
- <sup>21</sup> Mahzarin Banaji, "Implicit Attitudes can be Measured", in *The Nature of Remembering: Essays in Honor of Robert G. Crowder*, (Washington: American Psychological Association, 2001) ed. H.L. Roediger, J.S. Nairne, I. Neath, and A. Surprenant.
- <sup>22</sup> Marc Hauser, *Moral Minds*, 214.