

## **Moral Dumbfounding: When Intuition Finds No Reason**

Jonathan Haidt, Fredrik Björklund, and Scott Murphy  
University of Virginia

August 10, 2000

[Note: This is a temporary report of one study, based closely on the honors thesis of Scott Murphy. We ran a second study that manipulated cognitive load and found that load increased dumbfounding, but did not affect judgment. However, it is so time consuming to code all the videos that even now, in 2010, I still haven't fully analyzed that study for publication. And in social psychology one cannot simply publish a description of an interesting phenomenon, which is what this report is.]

### **Abstract**

Are moral judgments based on reason, or on intuition and emotion? Thirty participants were presented with a classic moral reasoning dilemma, and with four tasks that were designed to put intuition and reason into conflict. It was hypothesized that participants' judgments would be highly consistent with their reasoning on the moral reasoning dilemma, but that judgment would separate from reason and follow intuition in the other four tasks. This prediction was supported. In the four intuition stories (but not in the reasoning dilemma) judgment preceded reasoning, judgments were based more on gut feelings than on reasoning, and participants more frequently laughed and directly stated that they had no reasons to support their judgments. This phenomenon -- the stubborn and puzzled maintenance of a judgment without supporting reasons -- was dubbed "moral dumbfounding." The existence of moral dumbfounding calls into question models in which moral judgment is produced by moral reasoning. These findings are linked to other dual-process theories of cognition.

### Moral Dumbfounding: When Intuition Finds No Reason

How do we know what is right and what is wrong? On what is morality based? These questions are as old as philosophy itself. Plato (trans. 1973) held that the Form of the Good was directly apprehended through the study of philosophy. For Aristotle (trans. 1953), the good was not a mystical metaphysical unity, but rather a mixed bag of virtues. By habituating oneself to these virtues, one reaches *eudaimonia*, a kind of moral well-being. But both of these early philosophers agreed that the control of the passions by reason was essential to virtue and morality. For more than two thousand years after that, most philosophers have agreed with them.

It was not until the middle of the eighteenth century--during the "Age of Reason," no less -- that the dominance of reason in morality came under serious attack, in the writings of Scottish philosopher David Hume (1739/1969). Hume observed that "nothing is more usual in philosophy, and even in common life, than to talk of the combat of passion and reason, to give the preference to reason, and assert that men are only so far virtuous as they conform themselves to its dictates" (p. 460). He noted that reason was held to be eternal, invariable and divine, while passion was held to be blind, inconsistent, and deceitful. "In order to show the fallacy of all this philosophy," he continued, "I shall endeavor to prove, first, that reason alone can never be a motive to any action of the will; and, secondly, that it can never oppose passion in the direction of the will" (p. 460).

Hume did not fully succeed in his philosophical proofs of the impotence of reason. The present study, however, tests Hume's claims empirically. No study could show that reason can never oppose passion in the direction of the will, and indeed we think this hyperbolic claim is unlikely to be true. However we can investigate a class of moral dilemmas in which reason and passion conflict. If Hume is (generally) correct then passion will determine judgment and people will follow their feelings, even when they lack reasons to support those feelings. If Hume is incorrect then reasoning will precede judgment, and judgment will not be made without reasoning.

However before we can begin an empirical test of Humean psychology we must bring its terms up to date. Hume's most radical claim about human judgment was 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" (1739/1969, p.462). But what, in modern terms, are passions? In other passages from the same work Hume gives examples of passions such as anger, hope, fear, grief, joy, despair or security, suggesting that he means what we now call the emotions, broadly construed. Elsewhere in the work he discusses the emotions of "aversion" and "propensity", which motivate us to "avoid or embrace what will give us this uneasiness or satisfaction" (p.461). In modern terms he appears to be discussing a general approach-avoidance system (Davidson, 1992; Dollard and Miller, 1950). This approach-avoidance system is particularly important in the moral domain, giving us a "general appetite to good and aversion to evil" (p.464). As such Hume, like his fellow Scotsman Adam Smith (1759/1966), was proposing an innate moral sense (see Wilson, 1993, for a modern versions of the moral sense). Hume argued that this moral sense gives us certain "calm passions" which, because they do not cause as much "disorder to the soul"

as the emotions, are often mistaken for the products of reason. In modern terms these calm passions might be called intuitions, or, more popularly, “gut feelings”. In the present article we shall use the term “intuition” as the broadest modern term for what Hume meant by “passion”. Intuition should be understood to include a variety of automatic and uncontrollable cognitive processes, including emotional appraisals (Scherer, 1984) and automatic processes (Wegner & Bargh, 1998) which are largely outside the control of consciousness and independent of reasoning (see Bastick, 1982, for a review).

But what about the relationship between intuition (passion) and reason? Hume used the metaphor of master and slave, which we suspect will fail to resonate (or worse) with modern audiences. We can update this metaphor while still preserving Hume’s skepticism towards reason as follows: “reason is the press-secretary of the intuitions, and can pretend to no other office than that of ex-post facto spin doctor.” In modern political life the President makes his decisions first and then dispatches the press-secretary to justify and rationalize those decisions. The press secretary may have no access to the real causes of the President’s decision, and is therefore free to make up whatever argument will sound most convincing to the general public. Everyone knows that it serves no purpose to argue with the press secretary. Convincing her that her arguments are specious or that the President’s decisions are wrong will have no effect on the president’s decisions, since those decisions were not based on the press secretary’s arguments.

Several modern psychological theories have posited a similar ex-post facto role for reasoning. Nisbett and Wilson (1977) showed a variety of cases in which people’s behavior or judgment was influenced by factors outside of their awareness. Yet when asked to explain their behavior people promptly constructed plausible sounding explanations using implicit causal theories. Haidt, Koller and Dias (1993) observed a similar phenomenon when interviewing people about harmless violations of taboos, such as eating one’s (already dead) pet dog, or cleaning one’s toilet (in private) with one’s national flag. Participants often stated immediately and emphatically that the action was wrong, and then began searching for plausible reasons. Participants frequently tried to introduce an element of harm, for example by stating that eating dog meat would make a person sick, or by stating that a person would feel guilty after voluntarily using her flag as a rag. When the interviewer repeated the facts of the story (e.g., that the dog was thoroughly cooked so no germs were present), participants would often drop one argument and begin searching for another. It appeared that judgment and justification were two separate processes; the judgment came first, and then justification relied on “implicit moral theories” (paraphrasing Nisbett & Wilson, 1977), such as that moral violations have victims.

Shweder and Haidt (1994) drew on such observations to support a theory of “cognitive intuitionism” in which the human mind has been built to respond to certain moral goods. These goods appear to us as self-evident truths. They are not figured out or derived from first principles, although cultures have some leeway in making some of these goods more or less self-evident to their members (i.e., the goods of equality and autonomy may or may not trump the goods of chastity and piety).

Thus numerous theorists have suggested that moral judgment and moral reasoning may be two independent processes, in which moral judgment is like a form of perception (quick, non-

verbal, effortless, and utterly convincing) while moral reasoning is an ex post facto process, akin to explaining in words what one sees with one's eyes. A very useful language for discussing these two processes comes from Margolis (1987) who calls them "seeing-that" and "reasoning why". Margolis theorizes that the structure of the human brain cannot be radically different from its evolutionary ancestors. Our brains have been structured by millions of years of evolution for the function of pattern recognition, and our higher cognitive processes (such as reasoning) are only recent abilities being carried out by these same old structures. Although we may fancy that human cognition takes place exclusively using language and logic, the brain's structure and evolutionary history imply that most of our cognition involves lower or simpler processes of pattern matching.

Realizing that human cognition is based primarily on pattern recognition and similar primitive thought processes--what Margolis (1987) calls "P-cognition", for pattern cognition--clears up a number of mysteries of human cognition. For example, as Margolis points out, human cognition is notoriously bad at logic problems, e.g., the Wason 4 card task. Because the brain's structure is set up for P-cognition, and not logic, this is understandable.

Another seeming peculiarity in human cognition that is explained by P-cognition is the observations of the Gestalt psychologists. Gestaltists mostly studied visual perception, especially optical illusions. For example, Wertheimer (as cited in Wade, 1995, p. 137) often used a stimulus consisting of an image created of nothing but filled and unfilled dots; there are, in fact, nothing but filled and unfilled dots on the page, but what we perceive is the image. Gestaltists found that we do not passively perceive the world as it is, but rather actively organize it into wholes or patterns. Which is exactly what P-cognition predicts: we do not perceive the world as a field of visual data, and then "reason-why" particular configurations of this data represent distinct objects; rather, we quickly and intuitively "see-that" there are distinct objects, just as we "see-that" there is an image in Wertheimer's dot stimulus.

Margolis' (1987) theories are also in line with Nisbett and Wilson's (1977) work. Margolis holds that there are essentially two processes at work in human cognition, the evolutionarily old process of P-cognition, and the newly acquired process of verbal reasoning. The way these two processes work is that the P-cognition first provides a quick, pattern-matching "seeing-that," and then the reasoned, critical thinking process provides a post hoc "reasoning-why." This meshes perfectly with Nisbett and Wilson's observations that people often cannot accurately report on their mental processes; in fact, both Nisbett and Wilson and Margolis suggest that self-reports regarding mental processes generally are post hoc, and basically best guesses, based on the available information. While far from proving intuitionism, this does show that we are often mistaken when we claim our judgments are made based on facts about the world.

There are good theoretical grounds, therefore, for proposing that moral judgment might involve two separate processes: a quick, intuitive judgment ("seeing-that") followed by a slow, ex-post facto justification ("reasoning-why"). The present study tests this modern Humean proposal by placing participants in a situation in which the two processes are forcibly separated. We interviewed people about situations that were likely to produce strong intuitions that an

action was wrong, yet we engineered the situations to make it extremely difficult to find strong arguments to justify these intuitions. If Hume is right then people should cling to their intuitions, even in the absence of justification. If Hume is wrong then people should show a tight linkage between reasoning and judgment, and should not hold a judgment in the absence of reasons. We predicted that in these situations people would often make automatic, intuitive judgments, and then be surprised and speechless when their normally reliable “press-secretary” failed to find any reason to support the judgment.

To contrast with these situations we also gave participants a traditional moral judgment task: the Heinz dilemma (should Heinz steal a drug to save his dying wife?), from Lawrence Kohlberg (1969). Because this task requires participants to balance the interests of two people (the wife, versus the drugstore owner) we expected this task to be easy for the “press-secretary” to discuss. Because this moral dilemma has been so widely used in morality research we thought it important to determine whether the responses it elicits would be similar to the responses elicited by other kinds of judgment scenarios. We predicted that responses would not be similar, opening up the possibility that prior moral judgment research has tilted too heavily towards reasoning because it used a dilemma that was particularly easy to reason about.

## **Method**

### Participants.

Participants were 18 female and 13 male undergraduate students at the University of Virginia, who received credit towards an experimental participation requirement. One participant was 48 years old, and the rest ranged in age from 18 to 20. One participant refused to release the videotape made of her behavior, so the final sample consisted of 30 participants (17 female, 13 male).

### Materials.

Five tasks were used: one moral reasoning story, and four other stories designed to trigger intuitive judgments (see Appendix for full scripts). The “moral reasoning” story was taken from Kohlberg (1969). Commonly referred to as the “Heinz dilemma,” this story depicts a man (Heinz) who steals a drug to save the life of his dying wife. This story was chosen because it requires tradeoffs between competing interests and is therefore expected to trigger dispassionate moral reasoning. Furthermore this particular story has been the most widely used story in all of morality research, so it offers a clear anchor point for comparisons with other kinds of moral stories.

To contrast with the Heinz dilemma we used two “moral intuition” stories, written to be simultaneously harmless yet disgusting. One of these stories (Incest) depicts consensual incest between two adult siblings, and the other (Cannibal) depicts a woman cooking and eating a piece of flesh from a human cadaver donated for research to the medical school pathology lab at which she works. These stories were chosen because they were expected to cause the participant to quickly and intuitively “see-that” the act described was morally wrong. Yet since the stories were carefully written to be harmless, the participant would be prevented from finding the usual “reasoning-why” about harm that participants in Western cultures commonly use to justify moral

condemnation. For this reason we predicted that these two stories would produce different profiles of judgment when compared to the Heinz dilemma.

In addition we used two "non-moral intuition" tasks: Roach and Soul. Roach was taken from Rozin, Millman, and Nemeroff (1986). In this task the participant is asked to drink from a glass of juice both before and after a sterilized cockroach has been dipped into it. In the Soul task the participant is offered two dollars to sign a piece of paper and then rip it up; on the paper are the words "I, (participant's name), hereby sell my soul, after my death, to Scott Murphy [the experimenter], for the sum of two dollars." At the bottom of the page a note was printed that said: "this is not a legal or binding contract" (see Appendix). These tasks were designed to produce the same cognitive situation as the moral intuition tasks: a clear "seeing-that" the act was wrong or undesirable, coupled with a difficulty in finding "reasoning-why" to justify one's refusal. We included these two behavioral tasks to more clearly test Hume's predictions about the relationship between reason and action. Moral judgment tasks are only verbal actions, and we wanted to see what would happen when real behavior was called for.

#### Design and Procedure.

Participants were interviewed individually in a lab room equipped with a two-way mirror. Shelving and boxes covered all but a small portion of the mirror, obscuring it from view. A video camera was located behind the clear portion of the mirror, in an adjoining room, and a microphone was concealed in the ceiling above the participant's chair. To further convince the participant that he/she was not being videotaped the lab room also contained a large and conspicuous video camera on a tripod, visibly unplugged and pointed away from the participant.

After thanking the participant for taking part in the experiment, the experimenter told the participant that he/she would be presented with five "situations," in which the participant would be asked either to make a judgment or to do something, but that there was no right or wrong response. Participants were told that the experimenter would question their judgments or actions, and play "devil's advocate" by questioning their reasons. Participants were further told that they might find the stories or tasks objectionable, and that they could decline to participate in any given task, or even withdraw from the study entirely. After asking the participant to sign the informed consent form, the experimenter, gesturing vaguely in the direction of the unplugged video camera in the lab room, mentioned that the video camera would be used later in the study, but that, following the experiment, the participant would be given the option to refuse to allow his/her videotape to be analyzed.

The five stories/tasks were then presented in one of the two following orders, randomized within each gender separately, to counterbalance for order effects: Incest, Roach, Cannibal, Heinz, Soul; or Heinz, Cannibal, Roach, Incest, Soul. (The Soul task was always given last because pilot testing had shown it to be the "funniest" and therefore most mood-altering of the five tasks.) After each of the Heinz, Cannibal, and Incest stories were read, the participant was asked if what the depicted person or persons did was wrong; in the Roach and Soul tasks, the participant was asked to drink the "roached" juice, and to sign the "contract," respectively. The experimenter would then "argue" with the participant, non-aggressively, in an effort to undermine whatever reason the participant put forth in support of his/her judgment or action.

For example, following the Incest and Cannibal stories, if the participant responded that what the person or persons in the story did was wrong, the main counter argument was that no harm was done, and that the fact that an act is disgusting does not make it wrong. For the Heinz story, Kohlberg's (1969) "probe questions" were often drawn on; for example, if the participant responded that it was right for Heinz to steal the drug for his wife, he/she was asked if it would be just as right for Heinz to steal the drug for a stranger, or a pet animal that he loves. In the Roach task, if the participant refused to drink, the sterility of the cockroach was stressed, for example, by pointing out that it was cleaner than the juice. Lastly, in the Soul task, if the participant refused to sign, it was pointed out that he/she could immediately rip up the "contract," and that it was printed right on the "contract" that it was non-binding.

After the discussion that followed each task, the participant was asked to fill out a short questionnaire. The questionnaire asked the participant to respond on a Likert scale as to her level of confusion, irritation, and confidence in her judgment, and to what extent her judgment was based on reasoning or on a "gut feeling." After all five tasks had been completed, the participant was asked to fill out a final questionnaire that asked about a variety of demographic factors, including her own and her parents' political leanings and religious beliefs. Finally, the experimenter apologized for arguing with the participant, explained the hypothesis of the experiment, revealed that the session had been videotaped, and asked the participant whether she would grant permission for the tape to be analyzed. One participant refused.

## Results

Three kinds of data were obtained from each participant: (a) Likert scale self reports, after each task; (b) videotaped behavior on each task; and (c) demographic and personal information given by the participant after all tasks were completed. The major questions we asked of these data were: Is it possible to get people to say "I believe X is wrong, but I can't find a reason"? And if so, what other behaviors and features of judgment are observed? Furthermore, what kinds of tasks lead to the separation of reasoning and intuition? Is this separation equally likely whenever a person is argued with, or is it more likely in situations such as the Cannibal and Incest stories, in which strong intuitions contradict the absence of harmful consequences?

### Preliminary Analyses

Order effects. To determine whether it was appropriate to collapse across the two orderings of the five tasks independent samples t-tests were performed on all 180 Likert and videotape variables, for all five tasks (36 variables per task). This yielded a total of 180 separate t-tests, four of which yielded p values less than .05. Given the large number of tests and the small sample size per order we thought a Bonferroni adjustment of the alpha level was too stringent a correction, and we chose instead to look for patterns across the 5 tasks. Of the four "significant" tests, three involved variables for the Heinz story: When Heinz was given first, participants took longer to give their first evaluation, longer to give their first argument, and were more likely to change their minds. However given that the number of "significant" tests was fewer than would be expected by chance we concluded that results from the two orders were largely similar, and all subsequent analyses collapse across the two orders.

Gender effects. We performed the same sort of analysis to look for gender differences between the 17 female and 13 male participants. Out of 180 variables examined only 2 differed significantly at the .05 level. These two were that women had more conversational turns with laughter on the Roach task, and they took less time to utter their first word on the Roach task. An examination of those two variables on the other four tasks showed no consistent trend on time-to-first-word, however there was a non-significant trend in each of the other four tasks in which women laughed more than men, consistent with prior reports that women are more emotionally expressive than men (DePaulo, 1992). Because there were no other signs of gender differences, all subsequent analyses collapse across the data from men and women.

#### Differences between tasks

For each set of variables presented below a repeated measures MANOVA was performed across the five tasks to look for effects of task. The F and p values for each univariate test are reported in the right-most column of each table. Planned contrasts were performed between the Heinz task and each of the other four tasks, because we predicted that the Heinz task would be unique in encouraging analytical reasoning. Tasks that differed from Heinz at  $p < .05$  are marked with an asterisk. (Because the similarity between the four intuition tasks sometimes causes the overall F test for each variable to lose power we performed the planned contrasts even in cases where the F test did not show statistical significance at  $p < .05$ ).

Judgments and timings. Table 1 shows the basic judgments made by the 30 participants on all five tasks. As expected a strong majority said that it was OK for Heinz to steal the drug and it was wrong to eat human flesh or have consensual incest with one's sibling. The percentage of people willing to drink the roached juice was higher than we expected at 37%, but the percentage willing to sign the fake soul-selling contract was lower at 23%. After discussion with the interviewer the final judgments were slightly higher; that is, the interviewer did change some people's minds, in the direction for which he was playing devil's advocate, except that on the Heinz story the percentage endorsing Heinz' theft rose even though the interviewer was in most cases arguing against that position. The percentage of participants who changed their minds averaged 16%, and did not differ significantly across tasks.

The order of events in making judgments differed across the tasks. Table 1 shows that, on average, the first argument offered in the Heinz story preceded the first evaluation (judgment of right vs wrong) by 8.8 seconds. On the behavioral intuition tasks, however, the order was reversed, and participants generally first gave their evaluations and then offered reasons. The moral intuition tasks were split, with arguments preceding evaluations on the Cannibalism story by 5.9 seconds, but evaluations preceded arguments on the Incest story by 2.7 seconds.

Self-reports. : Table 2 shows the results of the self-reports that participants made on Likert scales after each task. There were significant differences due to task on five questions: 1) how sure participants felt about their judgments, with Heinz being high and the Cannibalism, Incest, and Soul stories significantly lower; 2) how confused participants felt, with Heinz being low and Cannibalism, Incest and Roach significantly higher; 3) how irritated participants felt while discussing the task; all tasks elicited low ratings of irritation, but Roach was lowest, followed by Heinz, and the Cannibalism story was significantly higher than Heinz; 4) the degree



to which participants said that their judgment was based on “careful reasoning about the facts and issues involved”, with Heinz being highest and the Incest, Roach, and Soul tasks significantly lower; 5) The difference between participants ratings of the degree to which they relied on reasoning versus “gut feelings”. Only on Heinz did participants say on average that their judgments were based more on reasoning than on gut feelings. In the other four tasks participants gave higher ratings to gut feelings than to reasoning, and this difference score was significantly different than Heinz for all but the Roach story.

Argument issues. Table 3 shows the means on variables coded from the videotapes pertaining to the kinds of arguments participants gave. There was a significant effect of task on how many arguments participants dropped (that is, repudiated, or at least stopped defending under cross-examination), with Roach and Heinz showing the fewest dropped arguments, while Cannibalism and Incest showed significantly more dropped arguments than did Heinz. There was a significant effect of task on the ratio of dropped to kept arguments, with Heinz showing the lowest ratio (.69), meaning that most arguments were kept, while both of the moral intuition stories showed significantly higher ratios, with approximately two arguments dropped for each one kept.

We observed a number of interesting response patterns in the videos, which we coded, and which are analyzed in the bottom half of Table 3. First, it often happened that participants made “unsupported declarations”, e.g., “It’s just wrong to do that!” or “That’s terrible!” They made the fewest such declarations in Heinz, and they made significantly more such declarations in the Incest story. Second, participants often directly stated that they were dumbfounded, i.e., they made a statement to the effect that they thought an action was wrong but they could not find the words to explain themselves. Participants made the fewest such statements in Heinz (only 2 such statements, from 2 participants), while they made significantly more such statements in the Incest (38 statements from 23 different participants), Cannibalism (24 from 11), and Soul stories (22 from 13). Third, participants often said “I don’t know,” sometimes several times in a row. There was a marginal effect of task in which participants said “I don’t know” least often in the Roach and Heinz tasks, and more often in the other three tasks. Fourth, we observed an interesting pattern in which participants would start giving an argument but as they were talking they realized that the argument was not going to work and they stopped in the middle of it, without any prompting from the experimenter. We called this pattern a “dead end”, and its frequency, while not significantly different across tasks, did show a pattern similar to the “I don’t know”, that is, lowest in the Roach and Heinz tasks, and higher in the Cannibal and Intuition tasks.

Non-verbal behavior. The behavioral intuition tasks generally took less time than the moral judgment tasks (122 seconds on average, for Roach and 292 seconds for Soul, compared to 379 for Incest, 399 for Cannibal, and 432 for Heinz). They therefore involved less time for non-verbal and quasi-verbal behaviors such as laughter and saying “um”. To correct for this time discrepancy we divided the total number of such behaviors per person per task by the number of minutes the task took. Table 4 therefore shows all values on a per-minute basis. Table 4 generally shows a split between the two behavioral tasks and the three moral judgment tasks. On

the two behavioral tasks participants were less likely to say “um”, more likely to laugh during each conversational turn, and more likely to touch their faces during each conversational turn (a potential sign of embarrassment according to Keltner, 1996) when contrasted with the Heinz story, which was generally similar to the two moral intuition stories. We also observed a small number of cases of a very unusual facial movement. It sometimes happened that participants would be giving a reason and then, while talking, they would pause or slow down, pull their eyebrows together and slightly down, yielding an expression that seems to observers to be the expression one would make if one was skeptical or doubting. We called these expressions “self-doubt faces” since they seemed to indicate that the participant doubted himself, even as he was talking. These faces were most frequent in the Incest story, although given the low number of times that such faces occurred we do not place much faith in this finding. We do, however, want to call it to the attention of future researchers. One additional measure showed no significant difference across tasks: the number of conversational turns in which the participant “fiddled with” or manipulated a pen with a prominent clicking mechanism (which we had placed on the table in the hopes that nervous or flustered participants would absent-mindedly manipulate it).

Summary of results differences between tasks. Putting together the results shown in Tables 1 through 4 tells the following story about the differences among the five tasks. Heinz conformed to the standard expectations of moral judgment research: arguments preceded evaluations, judgments were based more on reason than on gut feelings, most arguments were kept, people reported being fairly sure of their judgments, and they rarely said they could not explain their judgments. The picture on the two moral intuition tasks, however, was quite different. On those tasks the same participants reported being less sure and more confused, they reported relying on their gut feelings more than on their reasoning, they dropped most of the arguments they put forward, they frequently made unsupported declarations, and they frequently admitted that they could not find reasons for their judgments. The behavioral intuition tasks were generally more similar to the moral intuition tasks than they were to Heinz, except on the paralinguistic and non-verbal measures. Because these behavioral tasks were “funnier” than the judgment tasks, or perhaps because they required the participant to perform a real, self-relevant behavior, they elicited higher rates of laughter and face-touching.

Dumbfounding seems to occur when a strong intuition is left unsupported by articulable reasons. The clearest evidence of dumbfounding is that participants will often directly state that they know or believe something, but cannot find reasons to support their belief. These statements were coded as “Statements of Dumbfoundedness” in Table 3. They were made 38 times in response to the Incest story and 24 times in response to the cannibalism story, but only twice in response to the Heinz story.

### **Discussion**

Participants were often clearly dumbfounded by the moral intuition stories (Cannibal and Incest) and the nonmoral intuition tasks (Roach and Soul), while they were not dumbfounded by the moral reasoning task (Heinz). We also found some markers of dumbfounding. The most salient is that people who are dumbfounded will tell you so--they will say things like “I know it’s

wrong, but I just can't come up with a reason why." Participants who did so often also tended to report being more confused (in Roach and Cannibal) and to be relying more on "gut" than reason (in Incest and Soul); they also made more dead ends and unsupported declarations.

Nisbett and Wilson (1977) theorized that people generally do not have introspective access to the cognitive processes involved in making many judgments. Instead, people make ex post facto guesses as to what caused them to make judgments, based on the salient information they have available. Margolis (1987) theorized that, in general, human cognition usually involves a quick, intuitive "seeing-that," followed by a critical, ex post facto "reasoning-why," in order to explain why one came to the conclusion one did. This study provides support for the theory that moral judgments are often based on an intuitive, perhaps Humean feeling of rightness or wrongness, which is followed by "reasoning-why," based on the most salient features of the situation, in order to explain the judgment.

Kohlberg may have concluded that moral judgment was based on moral reasoning because the dilemmas he used, such as Heinz, had very salient fodder for post hoc "reasoning-why." In his dilemmas there were always questions of rights and harm (cf. Kohlberg, 1969). Had he used a broader sample of moral judgment tasks he might have come up with a different theory, one that gave greater prominence to moral emotions and the "seeing-that" of moral intuitions. (The tendency for psychologists to confuse a psychological phenomenon with the way they have chosen to study the phenomenon was called "the psychologists fallacy" by William James, 1890/1950).

The existence of moral dumbfoundedness – a state in which "seeing-that" conflicts with "reasoning-why" – is perhaps part of an emerging paradigm in psychology. This paradigm is illustrated by the work of Wegner (1994), who discusses the operations of "automatic" mental processes and "controlled" mental processes. Some mental processes, like "seeing-that," take place automatically, with no conscious effort. Other mental processes, like "reasoning-why," require effort and a conscious "controlling" of thought. Wegner (1994) found that these two processes come into conflict when one tries to suppress a particular thought, in a phenomenon he calls "ironic processes." As a result of one's efforts to suppress a thought via the controlling mental processes, ironically, one triggers the automatic mental process to produce the very thought that one is trying to suppress.

Perhaps the broadest conceptual basis for this paradigm of automatic and controlled processes can be found in Margolis' (1987) theory of P-cognition. From this perspective the automatic thought processes are our evolutionary inheritance of P-cognition. These are the same processes that lower animals use today, and that our evolutionary forbearers used throughout history, to navigate through the world. And perhaps the controlled processes are uniquely human cognitive abilities. One of these is reasoning, by means of which we are able to construct arguments, and attempt to explain or control our thought processes. It is no wonder that we are so poor at such controlled processes; the automatic processes have had a head start of several million years.

## References

- Aristotle. (1986). The Nichomachean Ethics (D. Ross, Trans.). New York: Oxford University Press (trans. 1953).
- Bastick, T. (1982). Intuition: How we think and act. Chichester, England: Wiley.
- Damasio, A. R. (1994). Descartes' Error. New York: Avon Books.
- Davidson, R. J. (1992). Emotion and affective style: Hemispheric substrates. Psychological Science, *3*, 39-43.
- DePaulo, B. M. (1992). Nonverbal behavior and self-presentation. Psychological Bulletin, *111*, 203-243.
- Donagan, A. (1977). The theory of morality. Chicago: Chicago University Press.
- Haidt, J., Koller, S. H., & Dias, M.G. (1993). Affect, culture, and morality, or is it wrong to eat your dog? Journal of Personality and Social Psychology, *65*, 613-628.
- Haidt, J., Rozin, P., McCauley, C., & Imada, S. (in press). Body, psyche and culture: Relationship between disgust and morality. In G. Misra, (Ed.), The cultural construction of social cognition. Sage Publications.
- Hume, D. (1969). A treatise of human nature. London: Penguin. (Original work published 1739 & 1740)
- James, W. (1950). The principles of psychology. New York: Dover. (Original work published 1890).
- Kant, I. (1981). Grounding for the metaphysics of morals (J. W. Ellington, Trans.). Indianapolis, IN: Hackett Publishing Co. (original work published 1785).
- Keltner, D., & Buswell, B. N. (1996). Evidence for the distinctness of embarrassment, shame, and guilt: A study of recalled antecedents and facial expressions of emotion. Cognition and Emotion, *10*, 155-171.
- Kohlberg, L. (1969). Stage and sequence: The cognitive-developmental approach to socialization. In D. A. Goslin (Ed.), Handbook of socialization theory and research. Chicago: Rand McNally.
- Kohlberg, L. (1971). From is to ought: How to commit the naturalistic fallacy and get away with it in the study of moral development. In T. Mischel (Ed.), Psychology and genetic epistemology (pp. 151-235). New York: Academic Press.
- Margolis, H. (1987). Patterns, thinking, and cognition. Chicago: University of Chicago Press.
- Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. Psychological Review, *84*, 231-259.
- Plato. (1987). The Republic and other works (B. Jowett, Trans.). Garden City, New York: Anchor Books (trans. 1973).
- Rozin, P., Millman, L., & Nemeroff, C. (1986). Operation of the laws of sympathetic magic in disgust and other domains. Journal of Personality and Social Psychology, *50*, 703-712.
- Scherer, K. R. (1984). On the nature and function of emotion: A component process approach. In K. R. Scherer & P. Ekman (Ed.), Approaches to emotion. (pp. 293-317). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Shweder, R. A., & Haidt, J. (1993). The future of moral psychology: Truth, intuition, and the pluralist way. Psychological Science, *4*, 360-365.
- Shweder, R. A., Mahapatra, M., & Miller, J. (1990). Culture and moral development. In J. Stigler, R. Shweder, & G. Herdt (Eds.), Cultural psychology. New York: Cambridge University Press. (Original work published 1987)
- Smith, A. (1966). The theory of moral sentiments. Trans.) New York: Kelly. (Original work published 1759).
- Wade, N. (1995). Psychologists in word and image. Cambridge, MA: MIT Press.
- Webster's third new international dictionary. (1994). Springfield, MA: Merriam-Webster.
- Wegner, D. M. (1994). Ironic processes of mental control. Psychological Review, *101*, 34-52.
- Wegner, D., & Bargh, J. (in press). Control and automaticity in social life. In D. Gilbert, S. T. Fiske, & G. Lindzey (Ed.), Handbook of social psychology (4th ed.). New York: McGraw Hill.
- Wilson, J. Q. (1993). The moral sense. New York: Free Press.

Table 1  
Basic Judgments

	Moral Reason	Moral Intuition		Behavioral Intuition		
	Heinz	Cann.	Incest	Roach	Soul	<u>F</u> and <u>p</u>
Initial judgment (%yes/OK)	80	13	20	37	23	n.a.
Final judgment (%yes/ok)	93	28	32	47	40	n.a.
% who changed	13	17	17	10	23	Chi-sq=2.44, n.s.
Seconds to 1st argument	11.2	9.6	12.6	6.2*	24.8*	16.72, p<.001
Seconds to 1st evaluation	20.0	15.5	9.9	3.9*	18.3	4.14, p=.01
Eval. precedes Arg. by:	-8.8	-5.9	2.7	2.3	6.5	3.03, p<.05

Note. % who changed includes changes in both directions, so is sometimes more than the difference between Initial and Final judgments. Changing judgment is a binary variable, so % changed is tested with a Friedman test.

\* = significantly different from Heinz at p<.05

Table 2  
Mean Self-ratings on Likert scales

	Moral Reason	Moral Intuition		Behavioral Intuition		
	Heinz	Cann.	Incest	Roach	Soul	<u>F</u> and <u>p</u>
How sure are you?	6.20	5.10*	5.37*	5.83	5.43*	2.34, p=.059
How much mind changed?	2.50	2.87	2.77	2.20	2.97	1.09, n.s.
How confused were you?	2.87	4.03*	4.00*	1.97*	3.43	8.77, p<.001
How irritated were you?	1.77	2.57*	2.20	1.43	2.00	5.40, p<.01
Judgment based on gut?	4.50	4.93	5.13	4.57	5.07	.97, n.s.
Judgment based on reason?	5.03	4.27	3.87*	3.90*	3.47*	3.67, p<.01
Gut minus reason:	-.53	.67*	1.27*	.67	1.60*	2.70, p<.05

Note. On the Likert ratings, 1=no/low, 7=yes/high

\* = significantly different from Heinz at p<.05

Table 3  
Means of Argument Variables

Variable	Moral Reason	Moral Intuition		Behavioral Intuition		F and p
	Heinz	Cann.	Incest	Roach	Soul	
Arguments dropped	2.9	6.4*	6.0*	2.7	3.60	8.40, $p < .001$
Arguments kept	4.2	3.2	3.2	3.1*	3.63	.90, n.s.
Ratio dropped/kept	.69	2.0*	1.87*	.87	.99	5.96, $p < .001$
Unsupported declarations	.8	1.9	2.4*	.1	1.00	3.03, $p < .05$
Statemnt/dumbfound'ness	.1	.8*	1.3*	.6	.7*	4.07, $p < .01$
I-don't-knows	1.4	2.3	2.4	.9*	2.03	2.21, $p = .077$
Dead-ends	.50	.83	.83	.20	.57	1.47, n.s.

Note. \* = significantly different from Heinz at  $p < .05$

Table 4  
Means of Paralinguistic and Non-Verbal Variables, Per Minute, Across Story/Tasks

Variable	Moral Reason	Moral Intuition		Behavioral Intuition		F and p
	Heinz	Cann.	Incest	Roach	Soul	
ums, uhs, hmms	1.98	1.70	1.94	.98*	1.25*	5.70, $p < .001$
turns with laughter	.55	.94*	.69	2.54*	1.62*	18.79, $p < .001$
turns with face touch	1.13	1.16	1.06	1.85*	1.77*	3.83, $p < .01$
doubt faces	.06	.05	.14	.01	.02	4.71, $p < .01$
turns with pen fiddle	.53	.80	.50	.88	.71	1.09, n.s.

Note. \* = significantly different from Heinz at  $p < .05$

## Appendix: The five situations:

## 1) The Heinz Dilemma

In Europe, a woman was near death from a very bad disease, a special kind of cancer. There was one drug that the doctors thought might save her. It was a form of radium for which a druggist was charging ten times what the drug cost him to make. The sick woman's husband, Heinz, went to everyone he knew to borrow the money, but he could only get together about half of what it cost. He told the druggist that his wife was dying, and asked him to sell it cheaper or let him pay later. But the druggist said, "No, I discovered the drug and I'm going to make money from it." So, Heinz got desperate and broke into the man's store to steal the drug for his wife. Was there anything wrong with what he did?

## 2) The Cannibalism Story

Jennifer works in a medical school pathology lab as a research assistant. The lab prepares human cadavers that are used to teach medical students about anatomy. The cadavers come from people who had donated their body to science for research. One night Jennifer is leaving the lab when she sees a body that is going to be discarded the next day. Jennifer was a vegetarian, for moral reasons. She thought it was wrong to kill animals for food. But then, when she saw a body about to be cremated, she thought it was irrational to waste perfectly edible meat. So she cut off a piece of flesh, and took it home and cooked it. The person had died recently of a heart attack, and she cooked the meat thoroughly, so there was no risk of disease. Is there anything wrong with what she did?

## 3) The Incest Story

Julie and Mark, who are brother and sister are traveling together in France. They are both on summer vacation from college. One night they are staying alone in a cabin near the beach. They decide that it would be interesting and fun if they tried making love. At very least it would be a new experience for each of them. Julie was already taking birth control pills, but Mark uses a condom too, just to be safe. They both enjoy it, but they decide not to do it again. They keep that night as a special secret between them, which makes them feel even closer to each other. So what do you think about this? Was it wrong for them to have sex?

## 4) The Roach Task:

*Experimenter asks:* Do you like apple juice?

*if "Yes":* Good.

*if "No":* OK, then, I have some water.

*Experimenter brings the appropriate beverage, a napkin, a cup, the roach container, and the tea ball to table.* OK, I have here a (can of apple juice/carton of spring water), which I'm going to pour into this glass [*pour it into glass*]. Would you be willing to take a sip of the juice/water? [*wait for S to take sip*]. OK, now I have here in this container some sterilized cockroaches. We bought some cockroaches from a laboratory supply company [*show box and label*]. The roaches were raised in a clean environment. But just to be certain, we sterilized the roach again in an autoclave, which heats everything so hot that no germs can survive. I'm going to dip this cockroach into the juice/water, like this. Now, would you take a sip of the juice/water?

5) The Soul Task:

I have a piece of paper here. If you agree to sign it, I'll give you two dollars, for real. If you sign it, you can then rip up the paper immediately, and keep the pieces yourself. So take a look at this [*hand S the "contract", which says:*].

I, \_\_\_\_\_,

hereby sell my soul, after my death,

to \_\_\_\_\_,

for the sum of \_\_\_\_\_.

\_\_\_\_\_

(signed)

Note: This form is part of a psychology experiment.

It is NOT a legal or binding contract, in any way.