

Handout #3: Haidt's Two Systems View

1. Characterizing the Opponent

The First Characterization: Haidt says that according to “rationalist models... moral judgments are caused by moral reasoning” (p. 814).

1st Problem: Reasoning from what? Even if a rationalist theorist says that moral judgments are the conclusions of reasoning, they must say something about the **premises** from which we draw these conclusions .

2nd Problem: Rationalist intuitionists (e.g., Kant, Robert Audi) think **some** general moral truths can actually be **known** in a non-inferential manner. And all rationalist theorists that I know of allow that unjustified or **unwarranted** moral judgments or beliefs are often the product of something other than reasoning (e.g. disgust and indoctrination). So if Haidt's rationalist is supposed to think that **all** of our moral judgments are generated by reasoning, Haidt's rationalist believes what no actual rationalist believes. Haidt is arguing against a “straw man.”

But if Haidt is saying that **no** moral judgment is caused by moral reasoning, as is suggested by his claim that he is considering the hypothesis that “Moral reasoning does not cause moral judgment,” he is asserting something fairly obviously false. Indeed, the positive model Haidt endorses allows what Kahneman would call “system 2” effects on the intuitive judgments that are products of system 1. Haidt allows that the moral beliefs of philosophers and students of philosophy are often modified (and so “causally influenced”) by various forms of reasoning.

The Second Characterization: When characterizing his alternative to rationalist models Haidt claims, “moral reasoning is usually a post hoc construction, generated after a judgment has been reached.” The use of “usually” here suggests that rationalists are committed only to the claim that moral judgments are usually the products of moral reasoning; that moral judgments are not typically intuitive judgments for which grounds are then sought.

Problem: Again, I know of no philosophical rationalist who makes this claim about the relative frequency of inferentially derived and non-inferential moral beliefs. As far as I can tell, this claim has nothing to do with “moral rationalism” as philosophers like Kant understand that position. Haidt attributes the view that moral judgments are usually the products of reasoning to Piaget, Kohlberg, and Turiel (psychologists in the Kantian tradition); but he doesn't provide quotes to substantiate these attributions.

2. Haidt's Social Intuitionism

Haidt's Intuitionism: The vast majority of our moral judgments (about particular cases) are generated by cognitive impressions or **intuitions**: quick, automatic evaluations (directly generated by our assessment of the non-moral facts of the case).

Haidt's Socialism: When reasoning does affect our moral judgments, it almost always does so via social and cultural interactions. [For example, the non-intuitive moral beliefs of a typical individual won't have been generated by reasoning done by that individual alone, but will owe their existence to the arguments and reasoning of theologians, politicians, and philosophers who are engaged in public debate.]

Question: Can one be a moral rationalist in the traditional sense **and** embrace social intuitionism?

Answer: Yes. Again, I think Haidt is "strawmaning" his opponents here.

3. Moral Dumbfounding

The Julie and Mark case: Incest without harm or injustice (Haidt, Bjorklund & Murphy, 2000).

Haidt's claims: Subjects have a cognitive intuition of the wrongness of the incestuous act directly generated by their knowledge of the facts of the case. Haidt calls the emotion generating judgments of the immorality of Julie and Mark's behavior "**revulsion** at the thought of incest." (p. 814). This intuition/emotional reaction **directly causes** subjects' judgments or their belief in the immorality of what's been done.

The Data: The case provides no good (non-theological) justification for the judgment that Mark and Julie are acting immorally by engaging in incest. But Haidt's subjects nevertheless try to invent justifications for the judgment even when they conflict with the (value-neutral) facts of the case as presented.

Haidt calls this phenomenon "moral dumbfounding."

Question: Must we experience revulsion to draw the judgment in question? Is revulsion correctly described as an intuition? In studying our reactions to incest, has Haidt selected a typical case of moral judgment? Do we also judge slavery and theft immoral because they fill us with revulsion? How can Haidt ground a judgment about frequencies (i.e. the claim that moral judgments are **typically** intuition-generated) from the observation (if it is that) that in a certain narrow range of experiments (having to do with incest, having sex with chickens, defiling the flag, etc.) moral judgments are generated by emotions like disgust or revulsion? Is it possible that we have good reasons to think incest is immoral but that we are reluctant to abandon the judgment in contrived cases engineered so that these reasons don't apply?

4. Haidt's Doubts About the "Causality of Reasoning in Moral Judgment"

- (a) Moral judgments are more often generated by intuition than reasoning.
- (b) The reasoning that does affect moral judgment is often "motivated" (i.e. inappropriately affected by the hopes, interests and emotions of the reasoner).
- *(c) It **seems** to us as though we reason our way to our moral convictions, but we **really** just invent reasons to justify moral beliefs that are in fact directly grounded in intuition.
- (d) Emotions have more of an effect in determining whether someone acts morally or immorally than does reasoning.

Question: Reflect on your experience in an ethics class. Is (c) true?

5. Haidt on the History of Moral Psychology

Haidt's interpretations of Hume and Kant are fairly accurate. But he overstates his case when he asserts, "Kant has had a larger impact than Hume on modern moral philosophers" citing Hare and Rawls as evidence (816). Hume's moral theory gave birth to Bentham's, and Bentham's gave birth to Mill's. We'd need to do a poll but I suspect that broadly consequentialist theories are at least as popular as deontic theories. Hume and Kant are at least tied as far as influence on contemporary ethics.

6. Kohlberg's Heinz Dilemma

The Dilemma: A woman was near death from a special kind of cancer. There was one drug that the doctors thought might save her. It was a form of radium that a druggist in the same town had recently discovered. The drug was expensive to make, but the druggist was charging ten times what the drug cost him to produce. He paid \$200 for the radium and charged \$2,000 for a small dose of the drug. The sick woman's husband, Heinz, went to everyone he knew to borrow the money, but he could only get together about \$1,000 which is 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.

Should Heinz have broken into the store to steal the drug for his wife? Why or why not? (Kohlberg, 1981)

Kohlberg doesn't focus on whether subjects say Heinz should or should not have stolen the drug. Instead, he focuses on the justifications subjects offer. This is supposed to reveal the kind of normative reasoning they use to arrive at their answer where normative reasoning normally develops from childhood to maturity in six stages:

Stage one (obedience): Heinz should not steal the medicine because he will consequently be put in prison which will mean he is a bad person. Or: Heinz should steal the medicine because it is only worth \$200 and not how much the druggist wanted for it; Heinz had even offered to pay for it and was not stealing anything else.

Stage two (self-interest): Heinz should steal the medicine because he will be much happier if he saves his wife, even if he will have to serve a prison sentence. Or: Heinz should not steal the medicine because prison is an awful place, and he would more likely languish in a jail cell than over his wife's death.

Stage three (conformity): Heinz should steal the medicine because his wife expects it; he wants to be a good husband. Or: Heinz should not steal the drug because stealing is bad and he is not a criminal; he has tried to do everything he can without breaking the law, you cannot blame him.

Stage four (law-and-order): Heinz should not steal the medicine because the law prohibits stealing, making it illegal. Or: Heinz should steal the drug for his wife but also take the prescribed punishment for the crime as well as paying the druggist what he is owed. Criminals cannot just run around without regard for the law; actions have consequences.

Stage five (human rights): Heinz should steal the medicine because everyone has a right to choose life, regardless of the law. Or: Heinz should not steal the medicine because the scientist has a right to fair compensation. Even if his wife is sick, it does not make his actions right.

Stage six (universal human ethics): Heinz should steal the medicine, because saving a human life is a more fundamental value than the property rights of another person. Or: Heinz should not steal the medicine, because others may need the medicine just as badly, and their lives are equally significant. [I've taken this analysis of the Heinz dilemma from http://en.wikipedia.org/wiki/Kohlberg's_stages_of_moral_development which you can consult for a further elaboration of Kohlberg's views and bibliographical references.)

Clearly subjects **infer** that Heinz should or should not have stolen the drug from their knowledge of the facts of the case. The question is whether there are **further normative premises** they use to derive their verdict.

Haidt's critique: Subjects make an intuitive judgment as to whether Heinz acted properly (where this judgment is grounded in a possible unconscious emotional or cognitive impression). The reasons they provide to justify their judgments are post hoc. To give Haidt his due: I think this is an interesting and important critique of Kohlberg's methodology.

Prior Critiques of Kohlberg: (1) Carolyn Edwards, "Societal Complexity and Moral Development" (1975); Carol Gilligan, *In a Different Voice* (1982). Demonstrated that Kohlberg's stages were biased toward western men. For example, Edwards tested moral leaders (e.g. judges and trusted arbitrators of tribal disputes) in Africa and they came out as comparable in moral development to American teenagers. (2) Turiel (1983) showed that young children don't think the wrongfulness of pulling hair etc is authority dependent contrary to Kohlberg's claims about stage 2 (cf. Damon, 1977 and Laupa & Turiel, 1986). (3) Hart and Fegley (1995) and Colby and Damon (1992) compared very

pro-social (helpful, engaged) people with those who were not and found no difference in the **Defining Issues Test (DIT)** score—a written exam (using a Likert scale) developed by James Rest as an alternative to Kohlberg’s interviewing technique.

Something to Consider: Mightn’t Haidt move too quickly from a rejection of Kohlberg’s model of moral judgment to a rejection of all broadly “rationalist” moral psychologies? Mightn’t Haidt move too quickly from a rejection of Kohlberg’s theory to the claim that his (i.e. Haidt’s) social intuitionism has been vindicated or supported?

7. Contemporary Descendants of Kohlberg: Eliot Turiel’s Social Interactionist Model

We’ve already discussed Turiel and Nucci’s work on **the moral/conventional distinction**: young children will say that if the teacher said it was ok to chew gum in class then it would be ok, but if the teacher said it was ok to pull someone’s hair in class it would still be wrong. Young Amish children said that if God decided that it was ok to work on the Sabbath then it would be ok to work on the Sabbath but if God said it was ok to hit people it would still be wrong to hit people.

Haidt interprets Turiel as arguing that people judge acts immoral when and only when they harm, involve injustice or violate rights. Acts that don’t have these features may be judged “inappropriate” or contrary to non-moral rules or conventions, but they won’t be judged immoral. Against Social Interactionism (so interpreted) Haidt, Koller and Dias (1993) found cases in which people continue to judge immoral eating a dead pet, having intercourse with a chicken, and cleaning a toilet bowl with the national flag even though these acts fail to produce harm, injustice or violate rights.

Question: Is Haidt just being uncharitable here? Did Turiel really claim that **no one** believes in victimless crimes? Classic liberalism involves the claim that people **should not** be condemned for harmless acts that violate no one’s rights. But liberals only need to make this argument because they rightly judge that many people do believe in the immorality of acts that are not harmful (e.g. homosexuality).

Haidt on the Evaluation of Abortion: Turiel argues that people’s judgments concerning the morality of abortion are grounded in judgments about its harmfulness and their judgments about when life begins (or when the fetus becomes a person). Haidt argues against this view as he thinks we have a “gut reaction” to abortion that is either positive or negative, and then invent claims about the beginning of life to justify the judgment this gut reaction causes us to make. But his case for this criticism is grounded entirely in his experiments on judgments regarding the morality of incest, having sex with chickens and the like.

Question: How good is this inference? Might judgments of the immorality of sibling incest be most often directly generated by revulsion whereas judgments of the immorality of abortion are most often grounded in the prior judgment that a child’s life begins early

in fetal development and the prior judgment that children all have a right to life? (We here put to the side the role that religious beliefs might play in both cases.)

8. Defining the Social Intuitionist View

Haidt, “The central claim of the social intuitionist model is that moral judgment is caused by quick moral intuitions and is followed (when needed) by slow, ex post facto moral reasoning” (p. 817). To clarify this claim he seeks to provide “definitions” of “moral judgment,” “moral reasoning,” and moral “intuition.”

(1) Haidt’s definition of moral judgment: “evaluations [good vs. bad] of the actions or character of a person that are made with respect to a set of virtues held to be obligatory by a culture or subculture” (817).

Questions: How accurate is this definition? What is it to form an evaluation “with respect to” a set of virtues? Notice that Haidt defines “moral” (as it applies to kinds of judgment) via “value” (i.e. representing things as good or bad) and “virtue.” Is this circular? And why does he include these concepts but not the other important class of morally relevant concepts: “ought” and “right”? Is it analytic or definitional that moral judgments involve some reference to value, virtue and vice but not rights and matters of justice?

(2) Haidt’s definition of reasoning: “a conscious mental activity that consists of transforming given information about people in order to reach a moral judgment. To say that the moral reasoning is a conscious process means that the process is intentional, effortful and controllable and that the reasoner is aware of what is going on” (p. 818). He cites Bargh (1994) and Galotti (1989) in defense of this characterization.

Problem: Isn’t there such a thing as unconscious (or inarticulate) reasoning? Is reasoning invariably effortful? If I wake up and see water on the ground, mightn;t I effortlessly infer that it rained the other night without consciously entertaining the intermediate premise that it’s having rained provides the best explanation of the water I now see? Might Haidt’s mistaken definition of reasoning as invariably conscious and effortful have led him to mischaracterize various alternatives to his view? Might it lead him to commit some substantive error? **What if moral judgment is often the product of relatively automatic, non-conscious (or tacit) forms of reasoning?**

(3) Haidt’s definition of moral intuition: “the sudden appearance in consciousness of a moral judgment, including an affective valence (good-bad, like-dislike), without any conscious awareness of having gone through steps of searching, weighing evidence, or inferring a conclusion...One sees or hears about a social event and one instantly feels approval or disapproval” (p. 818).

Notice that Haidt’s moral intuitions are only one form of Kahneman’s cognitive impressions as not all impressions need be emotional.

Criticism: We need to distinguish: (a) the claim that moral judgments are always, or almost always or typically not based in reasoning (beyond the reasoning necessary to process the value-neutral facts of the case as given by the experimenter's description of it) from, (b) the claim that moral judgments are always or almost always or typically grounded in or in part constituted by emotional reactions.

* It's possible (indeed, I think it is likely) that both (a) and (b) are true, but they are two entirely distinct claims.

Question: If someone draws a conclusion about the immorality of an act directly from her knowledge of the value-neutral facts of the case is her judgment not "intuitive"? A positive answer yields a weird taxonomy. Instead we should say that intuitive moral judgments are those **directly** drawn from a value-neutral assessment of the facts of the case, where "direct" means there is no process of reasoning intervening between understanding the description and performing the evaluation. That is, it is compatible with the judgment being directly grounded in one's value-neutral knowledge of the case that there are intervening emotions and "cognitive impressions" of the sort posited by Kahneman. Haidt's view is then best phrased as the view that moral judgments are both intuitive and emotionally grounded.

9. Elaborating the Social Intuitionist Model

(1) Moral judgments appear in consciousness automatically and effortlessly as the result of moral intuitions.

A recurring question: Must there always be moral intuitions as distinct from the moral judgments subjects formulate directly on the basis of their knowledge of the value-neutral facts of the case?

Another question: Which further conditions (if any) must an emotion the subject feels upon hearing the case meet if that emotion is to be considered a moral intuition?

(2) Moral reasoning only (or typically) occurs post hoc in an effort to justify intuitive moral judgments when they are challenged.

(3) "Reasoned persuasion works not by providing logically compelling arguments but by triggering new affectively valenced intuitions in the listener" (p. 819).

(4) Expression of moral judgments also works to change the views of others through pressures to conform.

Haidt says, "These four links form the core of the social intuitionist model. The core of the model gives moral reasoning a causal role in moral judgment but only when reasoning runs through other people" (p. 819). But he then softens this claim to allow that "philosophers" are an exception as they alone can reason on their own to moral judgments.

(5) People (mainly philosophers) can alter (or correct) their intuitions with reasoning, but in these cases they will have a “dual attitude” (citing Wilson, Lindsey and Schooler, 2000) as the corrected intuition will persist and affect their non-verbal behavior and cognition (i.e. it will persist as an implicit attitude).

Comment: Something like this is intuitive when the judgment is grounded in disgust as you may still feel disgust even after you realize that disgust is not a reliable indicator of moral status and so withdraw the judgment that the disgusting act is immoral. (Note, though, that intuitively speaking it's the **disgust** the persists despite the change in judgment, not the initial judgment in the immorality of the act one finds disgusting.) But why are we supposed to conclude that **all** or even **most** moral judgments are grounded in strong emotional reactions of this kind. If they are not, why should we think the cognitive impressions must remain after they've been corrected?

(6) We can modify our intuitions in a non-social or individualistic manner by imaginatively taking on the perspectives of others. This is a simulation of the discursive method by which such revision is more commonly affected.