



HÁSKÓLI ÍSLANDS

Hugvísindasvið

Judgments, Motives and Akrasia

Ritgerð til B.A.-prófs

Gunnar Júlíus Guðmundsson

Nóvember 2010

Háskóli Íslands

Hugvísindasvið

Heimspeki

Judgments, Motives and Akrasia

Ritgerð til B.A.-prófs

Gunnar Júlíus Guðmundsson

Kt.: 2209842819

Leiðbeinandi: Mikael M Karlsson

Nóvember 2010

Fyrir Mömmu

Acknowledgments

While writing this paper I was fortunate to be able to converse and consult with two philosophers, without whom this paper would not have been possible, Mikael M. Karlsson, Professor at the University of Iceland and Nigel Nigel Dower, Professor Emeritus at the University of Aberdeen, who both provided valuable criticism and guidance. Both philosophers have contributed significantly to my philosophical development through lectures, written text and conversations.

I would especially like to thank Professor Karlsson who has been a role model throughout my undergraduate studies.

I also want to thank my loving wife Alana Michelle Odegard and mother-in-law Pat Odegard, who proofread my thesis drafts, provided me with encouragement over the course of this project and put up with me during my academic studies in general.

In this paper I wish to do two things. First, I seek to explain the nature of the strong internalist theories about motivation I call in the paper *commitment* theories. Second, I wish to develop and argue for my *two-part theory* of moral judgments.

Commitment theories are theories about moral judgments that make acting against them impossible in the absence of other, overriding moral judgments. I recognize three distinct approaches to achieve this commitment that I call motivational, epistemological and combinational commitments. From this discussion I then proceed to talk about the well-known phenomenon of *akrasia* (weakness of will, or acting against one's better judgment) that seems to contradict the commitment theories. I discuss two types of answers to this problem: one I call the Socratic denial and the other judgment-disconnection approach. Although I do recognize that both types of answers solve the apparent contradiction I argue that neither of them treat *akrasia* adequately. I then introduce and develop my *two-part theory* of moral judgments. According to this theory, moral judgments are combinations of two sorts of judgments: first-order judgments that are motivational and second-order judgments that are evaluative of first-order judgments. *Akrasia* is then explained by a mismatch between the two types of judgments. The motivational judgment causes the agent to do something that his second-order evaluative judgment then condemns.

Table of content

I Introduction.....	2
II Moral Judgments	4
III Commitment Theories	6
IV Epistemological and Motivational Commitment.....	8
V Combination Theories	10
VI Akrasia	12
VII The Socratic Denial.....	14
VIII Judgment-Disconnection Approach.....	16
IX A Two-Part Theory of Moral Judgments	18
X Four Connections.....	20
XI The Two-Part Theory and Indeterminacy	22
XII Conclusion	23
Works Cited.....	24

Judgments, Motives and Akrasia

*We must examine it carefully for we are discussing no trivial matter,
but how we ought to live.*

-Plato, Republic, 352d

I

Introduction

There has been disagreement about the nature of moral judgments and moral motivations since the days of ancient Greece, and this disagreement has continued right into the current literature of meta-ethics. Of particular interest is the connection between these two concepts: how do moral judgments link up to moral motivation? In the modern debate, philosophers have lined themselves under two banners: the *internalists*, who think that moral judgments are necessarily motivating, and the *externalists*, who think that the motivation has to come from somewhere else. In this paper I will examine accounts of moral judgments that seem to make acting against them impossible in the absence of other, overriding moral judgments. I will refer to these *strong internalist* theories about moral judgments as *commitment theories*, since, according to them, the agent is necessarily committed to perform whatever acts are required of him by the moral judgments. This treatment of judgments has its roots in the teachings of Socrates and carried on, in its strongest form, into the *non-descriptivist* theories that dominated analytic moral philosophy in the latter part of the twentieth century.

A recognized weakness of commitment theories has been the fact that we seem frequently to act against our moral judgments. This is the phenomenon of *akrasia* (weakness of will, or acting against one's better judgment). I will argue that examples of *akrasia* do not constitute decisive counter-examples to commitment theories, since the commitment theorist can convincingly argue that such examples have only the *appearance* of *akrasia*. I will call this strategy (the rejection of purported examples of *akrasia* as not being genuine) the *Socratic denial*, and I will discuss this type of strategy in section VII below.

I accept the Socratic denial as a valid response to the problem of *akrasia*, but I will argue that, by itself, it is deficient, since it does not “save the phenomenon” of *akrasia*. Theories in this area must be careful to account for all relevant phenomena. I will discuss two attempts to save the phenomenon of *akrasia* by what I call *the disconnection approach*, treated in the section below that discusses Aristotle and Allan Gibbard. I will argue that these attempts to rescue the phenomenon through disconnection of moral judgment from action do a good job of capturing our common-sense idea of *akrasia*. However I will reject this approach, since it fails to capture instances of what I refer to as *cool akrasia*. By “cool *akrasia*” I have in mind instances where the agent is in good emotional balance and has plenty of time to think before he makes his moral decision but refrains, nevertheless, from acting in accordance with it.

In the light of these discussions I will formulate and argue for a *two-part theory of moral judgments* that analyzes moral judgments in two distinct components: first-order moral judgments that are necessarily motivating (where I follow Socrates and Hare and subscribe to the Socratic denial), and second-order moral judgments, which are judgments about judgments that account for the feelings the agent has about the first-order judgments that move him to action. In this way I propose to account for the phenomenon of *akrasia* as recognition of one’s cognitive failures, in the spirit of Aristotle and Gibbard.

II

Moral Judgments

Moral judgments, along with judgments of prudence, have the appearance of being motivationally self-contained. “Why should I be prudent?” or “Why should I be moral?” are questions that are only asked in philosophical contexts. In non-philosophical discourse the motive to act morally and prudently is simply assumed. In the case of a judgment of prudence (e.g., “I am getting out of shape, I should start jogging in the morning”) we judge that something should be done because it is in our self-interest to do so. After having established that something is in our interest to do, we do not look any further for a reason to motivate us. If a question is raised about a prudential action, for example “Why do you jog?”, the answer will most often be a listing of the benefits of jogging for the human body. If at this stage a curious questioner raises the further question of why we want these benefits for our bodies, the answer, if it is not from a meta-ethicist, is likely to be just a repetition of all the benefits of jogging.

Similarly, a moral judgment such as “stealing is wrong” seemingly incorporates a motive not to steal. When judging that something is wrong, we do not look further for a motive to comply: it is sufficient that the act in question (e.g., stealing) is wrong. In *My Father Bertrand Russell*, Katherine Tait reminisces about a conversation she had with her utilitarian father when she was young that illustrates this point very well:

“I don’t want to! Why should I?”

“Because more people will be happier if you do than if you don’t.”

“So what, I don’t care about other people.”

“You should.”

“But why?”

“Because more people will be happier if you do than if you don’t.”¹

In contrast with prudence, however, the motive behind the moral judgment has a categorical, or overriding, nature that can be formulated as, “I *ought* not to X no

¹ Tait, Katherine. *My Father Bertrand Russell*. (Thoemmes Press, 1996), 184-185

matter what other non-moral considerations I might have". This categorical nature of moral judgments becomes apparent if we contradict it. Consider: "Stealing is wrong, unless you can get away with it". If someone were to utter this statement we would no longer recognize it as a moral judgment. We would suspect the speaker to have intended to do "wrong" in some figurative sense, perhaps as a statement about the sociological fact that stealing is frowned upon by society. The statement will then merely be analogous to the statement: "Driving faster than the speeding limit gets you a speeding ticket, but only if you are caught" which is hardly a moral judgment.

Moral judgments can, however, be overridden by other moral considerations. If we substituted the non-moral consideration with a moral one, "Stealing is wrong, unless it saves lives," we would still allude to its moral status.

Moral and prudential judgments include a reference to certain motives that the agent holds, but this reference to motivations is merely on the linguistic or conceptual level. Although a moral judgment has been uttered, we have no guarantee that the utterance actually reflects a motive of the agent. She could be misleading us, for example. If we knew that the agent sincerely accepted a certain moral judgment, a judgment that is conceptually tied to her categorical motivation, we would have an insight into the agent's state of mind and would be able to partially predict her future actions. However, this is not the case with judgments of prudence, since they can accurately be formulated as "if-then" sentences (e.g., "If I exercise regularly then I will feel better"), and if-then sentences are factual statements that can be accepted without any practical commitment if we are not already committed to the antecedent. The motivation for prudential judgments is therefore *external* to the judgment itself and will be outside of our scope in this paper.

III

Commitment Theories

There are three ways one can be committed to acting through acceptance of a moral judgment. First, it may be that the fact of what the moral judgment is about is such that the knowledge of it, in itself, provides a strong motive to act accordingly. This I will call *epistemological commitment*. It is the position of Plato and can be formulated as follows:

Making moral judgment X while possessing moral knowledge Z causes the agent to be motivated to do actions Y.

For Plato, moral knowledge is knowledge about the highest of all forms: the form of the good. The form is outside of the agent, objective and real. It is, however, still capable of providing its knower not only with information about what to do, but also of giving him an overriding reason to do so. This combination of moral realism and strong internalism is rare. Most moral realists of today treat moral facts like natural facts and take the externalist line (i.e., it depends on the agent if he is motivated by them or not). This is therefore the rarest type of commitment theory.

A second approach claims that sincere moral judgments are the result of the agent already having the motive to perform the relevant act. I call this *motivational commitment*, and it is the position of the non-descriptivist, R. M. Hare:

By being morally motivated to do action Y in situation C, the agent makes the moral judgment X.

According to this type of commitment theory, the agent already has the motivation that his moral judgment expresses and therefore the agent does not need any external reasons to be motivated. With the rise of the *non-cognitivist* theories, this approach to moral motivation has become very common. Non-cognitivist theories vary in the sort of mental state that is supposedly expressed in the moral judgment. Hare's position is that it expresses a motive for universal prescriptions. I will discuss this position in the next section.

Finally, we have the theory that the agent has certain properties, and the moral judgment has certain properties such that when they combine in the agent's making of a moral judgment, the agent is motivated to act in the manner prescribed in that judgment:

If the agent has property P, he will be motivated by moral judgment X to do Y.

This is the position of Socrates, who thinks that being moral is always in the best interest of the agent. Since he further thinks that it is in human nature to be motivated to do what we think is best for ourselves, he arrives at the necessity of being motivated by moral judgments. This type of commitment theory has strong supporters. One of the proponents of this approach is Kant, who thought that the rationality of morality was necessarily motivating for all rational beings. His approach is still popular today among Kantian ethicists.

I have now described the three types of commitment theories, *epistemological*, *motivational* and *combinational* and I will now examine these three types in turn.

IV

Epistemological and Motivational Commitment

In his *Republic*, Plato emphasizes the distinction between opinion and knowledge. Opinion is connected with faulty particulars, while knowledge takes the everlasting and unchanging forms as its object. True moral judgments are directed at the form of the good. John L. Mackie in *Inventing Right and Wrong* writes: “The form of the good is such that knowledge of it provides the knower with both a direction and an overriding motive [...] not because of any contingent fact that this person, or every person, is so constituted that he desires this end.”² Plato does not elaborate on this self-motivating property of the form of the good but proclaims that, through education, the minds of the rulers can be trained to fix only upon the forms, including the form of the good. Since the rulers have this knowledge, they can be entrusted with all the power of government, because, as knowers of the good, they will have no inclination to act wrongly.

In contrast to the epistemological account above, motivational commitments claim a causal relationship between moral motivations and moral judgments. These sorts of commitment theories are either *subjectivistic* (where the *sincere* moral judgment is taken to be a description of an agent’s mental state), or *expressivistic* (where the moral judgment is a non-descriptive expression of the mental state of the agent). Here I will discuss only the latter. In expressivism, the idea is that moral judgments do not describe anything and are therefore not capable of being true or false if truth conditions are understood to state something about the world. Moral judgments are therefore similar to exclamations like, “Boo, Manchester United!” and “Ouch!”

According to Hare’s *universal prescriptivism*, moral judgments express the state of mind wherein one wants to *universalize* a value by *prescribing* it to all agents. This universalization also applies to the person making the judgment. In *The Language of Morals*, Hare writes that the test for whether a person’s judgment, “I ought to do X”, is a moral judgment is the following: “Does he or does he not recognize that if he

² Mackie J.L. *Inventing Right and Wrong* (London: Penguin, 1990) 40

assents to the judgment, he must also assent to the command 'let me do X'".³ Hare also states "[i]t is a tautology to say that we cannot sincerely assent to a second-person command addressed to ourselves, and *at the same time* not perform it, if now is the occasion for performing it and it is in our (physical and psychological) power to do so".⁴ The sincere moral judgment is therefore an expression of the agent's motives to act according to the judgment. As a result, the agent is incapable of acting against his moral judgments if he has sincerely made the moral judgment.

Gibbard's position is similar to that of Hare. Gibbard agrees that moral judgments are expressions of universalizable prescriptions. However he disagrees with Hare's claim that one must necessarily act on them in order to hold them sincerely. Gibbard finds it sufficient that acting against the prescription would warrant sentiments of guilt and remorse in the agent. Therefore, moral judgments are "judgments of what moral feelings it is rational to have" and "to call something rational is to express one's acceptance of norms that permit it"⁵. I will give a similar treatment to what I call second-order moral judgments in the section where I introduce my two-part theory of moral judgments.

³ Hare, R. M. *The Language of Morals*. (Oxford: Clarendon Press, 2003), 168-169

⁴ Hare 2003, 20

⁵ Gibbard, Allan. *Wise Choices, Apt Feelings: A Theory of Normative Judgment*. (New York: Oxford University Press, 2002), 6-7

V

Combination Theories

Combination theories can be seen as a cross between epistemological and motivational theories about commitment. In the dialogue, *Protagoras*, Socrates asks Protagoras if he agrees with him that “knowledge is something noble and able to govern man, and that whoever learns what is good and what is bad will never be swayed by anything to act otherwise than as knowledge bids, and that intelligence is a sufficient succor for mankind.”⁶ This expresses what I have here called an epistemological commitment theory of moral motivation.

Socrates is a virtue ethicist, and virtue ethics makes no distinction between judgments of prudence and moral judgments, although they can (and do) make the distinction between prudential virtues and other sorts of virtues. They will still hold that it is beneficial for the agent to hold these non-prudential virtues. Virtue ethicists start with the question: “What is the good life?” and consider finding an answer to this question to be the goal of ethics. Living the good life is of course considered desirable and the judgment about what is the right thing to do in a given situation is therefore the same judgment about what is best for one’s self.

Socrates thinks that it is in our nature always to act in what we take to be in our own best interest: “... [N]o one willingly goes after evil or what he thinks to be evil; *it is not in human nature*, apparently, to do so”⁷ [emphasis mine]. For Socrates then, all of our actions are testimonies concerning what we think is best for us. The commitment can therefore be understood by the analogy of plugging in an electrical device. When a lamp is connected to the wall, it lights up. This is not to be explained only by the electrical outlet or by the lamp itself, but by the combination of the properties in the lamp and in the electrical wiring in the wall. In the same way, the judgment about what is the right thing to do, along with the agent who possesses human nature, will light up and the agent will act accordingly. This combination of properties in the agent and in the judgment therefore creates the necessity that

⁶ Plato. Edman, Irwin ed. *The Works of Plato*. (New York: Tudor Publishing Company, 1936) 352c

⁷ Plato 358 c-d

makes the agent committed to act on his judgments. If the agent holds that a particular action is best for him, he cannot fail to act on it.

I have now given examples of the three types of commitment theories that I listed in section IV. Although they approach moral motivation from very different, even opposite, angles, they all arrive at the common conclusion that it is impossible to act against one's moral judgments in the absence of other overriding moral judgments.

VI

Akasia

Acting against our better judgment is something that we seem to do all the time. We procrastinate, eat fatty foods and stay up later than we think is good for us. However, as section II makes clear, these are all examples of acting against judgments of prudence. But can we act against moral judgments? It was claimed in section II that moral judgments are judgments that claim priority over other considerations, or, at least, over all non-moral considerations. For example, I might say, "I know this chocolate bar is bad for me, but I like the taste so much that I am going to eat it anyway". Even if imprudent, I would not thereby be inconsistent. If I were, however, to utter, "I know that stealing this chocolate bar is wrong, but I like the taste so much I am going to steal it anyway," I would, in a way, contradict myself. How could I steal the chocolate bar if I believed that stealing it is something that I categorically ought not to do?

We have already seen that according to Plato, the mere possession of the knowledge that stealing a chocolate bar is wrong renders the agent incapable of doing so.

According to Socrates, thinking about stealing a chocolate bar is not in one's best interest and it is contrary to human nature to act against one's own perceived best interest and steal.

Also, according to Hare, thinking that stealing a chocolate bar is wrong is an expression of an already existing motivation to not steal.

On the surface, it may seem that such an act is impossible. However, we do seem to sometimes act against our moral judgments. For example, we may watch advertisements about children in need, but do nothing; drive polluting cars when we could take the bus; or lie in order to get ourselves out of trouble. In such cases, we willingly fail to do something that we judge to be something that we ought to do.

The contradiction is as follows:

- a. Necessarily, if one makes the moral judgment “I ought to X” then one intends to X
- b. I judge that in situation Y, I ought morally to X
- c. I judge that I am in situation Y
- d. I willingly fail to do X

Clearly the propositions a-d are contradictory since a-c require the opposite of d. Since d is, as we have seen, a well-known phenomenon of everyday life (not contested except perhaps on the philosophical level), the commitment theorist has to give an explanation of these *akratic* examples. To solve this contradiction, the commitment theorist has to deny at least one of the propositions b-d in order to not have to deny a, since denial of a is the rejection of the commitment thesis.

Since the contradiction derives from the commitment aspect of the theories now discussed, *akrasia* is a problem for all of them, and a solution to it should likewise salvage them all. In what I have called the *Socratic denial*, Socrates and Hare deny premise b, while Aristotle and Gibbard advocate what I call *disconnection approach* by rejecting c.

VII

The Socratic Denial

In the *Protagoras*, right after having proclaimed that “Whoever learns what is good and what is bad will never be swayed by anything to act otherwise than as knowledge bids,” Socrates considers the example of *akrasia*. He says “*most people* will not listen to you [Protagoras] and me, but say that many, while knowing what is best, refuse to perform it, though they have the power, and do other things instead”.⁸ The answer is quick to follow: “it is from the defect of knowledge that men err when they do err, in their choice of pleasures and pain that is in the choice of good and evil.”⁹

The Socratic denial thus consists in the claim that, in purported examples of *akrasia*, there is only the *appearance* of acting against one’s moral judgments. If someone lies to get herself out of trouble, she does not truly judge that lying is wrong, even if she declares that it is.

More than 2000 years later, Hare gives a similar response in the context of his universal prescriptivism. The reader will remember that universal prescriptivism is the idea that sincere moral judgments are prescriptions or commands that are supposed to apply to everyone, including the person making the judgment. Hare is a commitment theorist because he thinks that “[i]t is a tautology to say that we cannot sincerely assent to a second-person command addressed to ourselves and *at the same time* not perform it, if now is the occasion for performing it and it is in our (physical and psychological) power to do so.”¹⁰

In *Freedom and Reason*, Hare examines several different kinds of *akratic* examples. Here I will only look at the most relevant of them: The example of the agent who makes a moral judgment (and therefore prescribes himself an action) but cannot carry it through because of weakness of will. This type of *akrasia* Hare calls cases of “ought but can’t” and writes of them that the moral judgments of the acting agent can

⁸ Plato, 352d

⁹ Plato, 357d

¹⁰ Hare 2003, 20

either be withdrawn or down-graded: “[Such a judgment] no longer carries prescriptive force in the particular case, though it may continue to do so with regard to actions in similar circumstances”.¹¹ The prescription has therefore been changed from “Don’t do X in situations Y” to “Don’t do X in situations Y unless...” Note that in my account of moral judgments in section II, this would possibly no longer qualify as a moral judgment. Both moves on the behalf of the agent are examples of Socratic denial, since the judgment no longer applies to the case at hand.

The responses of both Socrates and Hare are adequate reactions to the *akratic* challenge. By denying the agent the knowledge of the moral judgment, or full acceptance in the case of Hare, the *akratic* examples no longer present a contradiction. The strategy of both philosophers is to deny the common human experience of thinking that something is really what is morally required of one’s self in a situation exactly like this and yet failing to act accordingly. The Socratic denial must therefore be considered inferior to a response that would be capable of solving the dilemma *and* recognizing this phenomenon as a genuine part of human experience.

¹¹ Hare, Richard Mervyn. *Freedom and Reason*. (London: Oxford University Press, 1970), 81

VIII

Judgment-Disconnection Approach

In his *Nicomachean Ethics*, Aristotle recognized this weakness of the Socratic denial and made it his starting point into the inquiry of *akrasia*: “The Socratic theory is evidently at variance with the facts of experience, and if ignorance be the cause of passion, i.e. of incontinence [*akrasia*] it is necessary to inquire into what is the nature of the ignorance”.¹² Aristotle makes a distinction between two types of *akrasia*, with or without deliberation: “There are two classes of incontinent people, and those who simply lose command of themselves are better than those who possess reason but do not abide by it, as they are not overcome by so violent an emotion, nor do they act without previous deliberation like the others.”¹³

In Aristotle’s syllogistic (his deductive logic), propositions b and c in section VI are what he calls the *major* and *minor* premises of an argument. The *akratic* person who does not deliberate never has the minor premise enter his stream of consciousness and thus acts for other reasons. It is not until later, when this *akratic* person reflects on his actions, that he feels the guilt and remorse of having acted wrongly. If he had deliberated prior to acting, he would have acted differently. In deliberative *akrasia*, the agent, after having deliberated, reaches the conclusion that he ought to do X. However, the agent experiences a conflict between acting out of his moral reasoning and his reasoning driven by pleasure or anger. The agent’s sentiments do not line up with his rationality, and he fails to act on his moral reasoning.

In both types of *akrasia*, there is a disconnect between a moral judgment and the action that ought to follow. In the non-deliberative case, the agent does not stop to think that a moral principle that he in fact accepts is applicable, while in the deliberative case, the judgment does not lead to the action because of a faulty alignment of sentiments. Aristotle thinks that knowledge of both the major and minor premises is sufficient for action and therefore thinks that the deliberate *akratic* person does not have that knowledge. Rather “we must suppose then that people in

¹² Aristotle. *The Nicomachean Ethics*. Trans J.E.C Weldon. (New York: Prometheus Books, 1987), Book 7 chapter 3

¹³ Aristotle, Book 7 chapter 3

a state of incontinence repeat phrases in the same way as actors on the stage.”¹⁴

He goes on to say, “in fact there is a sense in which he [the incontinent person] does possess knowledge and another sense in which he does not possess it, as e.g. in sleep, madness or intoxication. But this is the very condition of people who are under the influence of passion.”¹⁵

This method of disconnecting the moral judgment from the action seems only to work in *akratic* situations where the agent has to act or make the moral decision relatively quickly. We do however have cases of what I will call *cool akrasia*, where the agent has ample time to contemplate the action and can therefore look at his action in the light of the many moods he may be in. An example might be that, after having read about poverty relief, the agent forms the moral judgment “an affluent agent ought to share her wealth to reduce suffering” and that, since she is in a situation of relative overabundance, she ought to donate some money to charity. This thought grips the agent and she thinks about it constantly. It would take her only few minutes to donate by placing a phone call, but no matter what her emotional state, whether she is worrying about her bills or is in good balance, she never picks up the phone.

These sorts of examples cannot be explained away as cognitive hiccups, since the agent is presented with many different occasions for acting on his judgment, but still fails to act. The disconnection approach does therefore not comfortably capture all examples of *akrasia*.

¹⁴ Aristotle, Book 7 chapter 5

¹⁵ Aristotle, Book 7 chapter 5

IX

A Two-Part Theory of Moral Judgments

Practical judgments can be analyzed into two components: as a judgment about *what to do* and as a judgment about *how to feel* about the situation with which one is confronted. The judgments about what to do I call *first-order judgments*, and the judgments on how to feel about our actions I call *second-order judgments*. I will now discuss each in turn.

First-order judgments can be either descriptive (there is a hot teacup on the table), or evaluative (tea tastes good). Some of these first-order judgments will be motivating. I, for instance, would be motivated by the judgments mentioned above to have a sip of my tea. Whether the motivation will actually result in an action depends upon other judgments I hold. I might, for example, think that I have had enough caffeine for today, and that might keep me from having a sip of the tea. Of course it also depends upon one's personal preferences as to whether being in close proximity to hot tea moves a person in one way or the other. Moral judgments, however, as I discussed in section II, are motivationally self-contained. Therefore simply by the making a moral judgment we have the motivation to act on it. If we judge: "Giving to charity is morally good" we are moved to donate to charity. As I also mentioned in section II, moral judgments are categorical, claiming precedence over all non-moral judgments. So unlike the person who can resist the temptation to drink tea because he thinks he has had enough caffeine, the person who stands in front of the charity box has no such option. Either he donates, or he does not truly think that giving to charity is morally right in this sense, granted of course he has no overriding moral reason not to donate.

This is a rejection of *akrasia* and echoes the Socratic denial discussed in section VII. If an agent acts against a motivationally self-contained and categorical judgment, he cannot hold it himself. However, as I noted in section VIII, we often seem to act against categorical and motivationally self-contained judgments that we genuinely make. This is where the second-order judgments come into play.

Second-order judgments are evaluative and descriptive judgments about first-order judgments. A common form of evaluative second-order judgment is: “judgment X is good/bad” or similarly, “It is nice to like tea”, while descriptive second-order judgments take the form: “Judgment X is Y”, for example, “Judgments of prudence are judgments about what is good for oneself”.

Normally, second-order judgments about moral matters are linguistically identical to their first-order counterparts. This means that if we wanted to convey the idea that someone should feel bad about stealing or that he should not steal we could use exactly the same expression: “Stealing is wrong”. According to the two-part theory, moral judgments are composed of *both* first- and second-order judgments. They state simultaneously what we are to do, and how we are to feel about it.

X

Four Connections

First-order and second-order judgments can combine in four possible ways.

1. Rejection of both the first- and second-order moral judgment
2. Acceptance of first- and second-order moral judgments
3. Rejection of second-order moral judgment but acceptance of first-order moral judgment
4. Rejection of first-order moral judgment but acceptance of second-order moral judgment

The first pairing is a total negation of the moral judgment. The agent does not feel motivated not to X and does not think that he should feel motivated.

A negation of “X is wrong” is not “X is right,” but rather “It is not the case that X is wrong”, which does not have a built-in favorable attitude towards X. In this case the agent simply does not accept that “X is wrong” either as a motivating reason or as an attitude towards the motivation.

The second pairing is the opposite of the first. The agent is motivated by “X is wrong” and he experiences guilt if he fails to comply.

This is what we might call the common view of moral judgments. The agent, after having made the moral judgment, feels motivated to act accordingly and would feel guilt or remorse if he did not succeed.

A formulation of the third pairing in the example of “X is wrong” could be “I categorically should not X, but I would not feel bad if I did.”

These sorts of moral judgments are known in the literature as *supererogatory*. They concern morally praiseworthy deeds that are not considered obligatory. The agent avoids the morally bad X, but he would not feel guilty if he did not avoid it. That the

two-part theory is capable of covering this additional moral phenomenon must be counted in its favor.

The fourth pairing captures the subject of akrasia. In this case, for the agent, “X is wrong” is not motivating, but he thinks that he should find it motivating.

According to the theory, first-order moral judgments are categorically motivating. If the agent sincerely accepts them he will feel motivated to act on them. However, we cannot choose which judgments we find motivating and which we do not. Just as I cannot make myself believe that there is a Santa Claus in this room, I cannot help that I want this chocolate cookie. Here, I adopt Socratic denial and deny *akrasia* on the level of the first-order judgment. I therefore deny holding the moral judgment to the agent in this type of case. If moral judgment X does not motivate, it is not held as a first-order moral judgment.

Instead I propose that *akrasia* exists on the level of second-order moral judgments. It is the remorse or guilt about what actually motivates one to action where we wish that we were motivated by different first-order judgments. *Akrasia* in this way is therefore seen as a “disconnection” between first- and second-order moral judgments. The agent is not motivated not to X, but he thinks that he should be.

The two-part theory also has the benefit of being able to account for the *cool akrasia* I discussed in section VIII. Consider again the example about the agent who reads about poverty relief and judges that a person in her situation should send money, and although she thinks about it constantly, she never actually donates. According to the two-part theory, she has formed only the second-order judgment that one should feel motivated to donate money in this circumstance. However, she is not motivated by the first-order moral judgment and she therefore feels guilt and remorse.

This solution accounts for our intuitions about *akrasia* and the accompanying guilt and remorse and is therefore preferable to Socratic denial alone. Furthermore, the two-part theory is capable of accounting for all of the examples of *akrasia*, even those which I have here called *cool akrasia*.

XI

The Two-Part Theory and Indeterminacy

A result of the two-part theory is that the agent does not necessarily know whether he is making a first- or second-order moral judgment. After having made a moral judgment that applies to situations that the agent is unfamiliar with, the agent may not be capable of determining whether the judgment moves him when the occasion arises. This happens because first- and second-order moral judgments are normally linguistically identical and because we do not have a perfect understanding of our psyche. When the agent looks inward, he can feel that he truly accepts that X is the right thing to do in situations Y (situations that he has not been in before). However, unless he is at that moment in situation Y or has a good understanding of his own psychology, he will be at loss of whether he will act accordingly in that situation.

This indeterminacy of how the agent will act after having made a moral judgment brings the theory away from the commitment theories discussed above. Although the agent makes the moral judgment sincerely, we are not capable of predicting with certainty the actions of the agent when the relevant circumstance arises.

Predictability was a common element in the motivational, epistemic and combinational theories discussed above. However, the unpredictability that exists in the two-part theory does not arise because moral judgments are not necessarily motivating, but rather because the agent cannot determine which type of moral judgment it is that he holds.

XII

Conclusion

I have discussed the overriding property of moral judgments and the commitment theories that take these judgments at face value. In sections IV-V, I discussed *epistemological*, *motivational* and *combination-commitment* theories and their common recognition of the necessary efficaciousness of moral judgments. In section VI, I explained the phenomenon of *akrasia* and showed how, if accepted at face value, it creates a contradiction for the commitment theories. In sections VII and VIII, I looked at two proposed ways to solve the dilemma: Socratic denial (the denial of holding the moral judgment) and the disconnection approach (appealing to a mental hiccup that prevents the moral judgment from leading to the appropriate action). I found that both responses were sufficient to account for certain cases of *akrasia*, but that both failed to account for other cases recognized in our common experience. As Aristotle pointed out, Socratic denial rejects commonly recognized phenomena; but the disconnection approach fails to explain instances of *cool akrasia*. In section IX, I put forward a two-part theory of moral judgments where I divorced first-order, motivational judgments from second-order moral judgments that dictate our feelings about our actions. I proposed to adopt Socratic denial for first-order moral judgments: these, I maintain, are necessarily efficacious. I then used aspects of the disconnection approach to explain the emergence of *akrasia* from the disagreement between second-order and first-order judgments. In section X, I considered the possibilities for combining first- and second-order judgments and maintained that my theory seems to account for all of the possible attitudes the agent can have towards a moral judgment. Finally in section XI, I discussed the indeterminacy that is inherent in the two-part theory that is not present in the epistemological, motivational and combinational theories. In this way, I have attempted to solve both the *akratic* contradiction for commitment theories and to save the phenomenon of *akrasia*.

Works Cited

Aristotle. *The Nicomachean Ethics*. Trans. J.E.C Welldon. New York: Prometheus Books, 1987.

Gibbard, Allan. *Wise Choices, Apt Feelings: A Theory of Normative Judgment*. New York: Oxford University Press Inc, 2002.

Hare, R. M.. *Freedom and Reason*. London: Oxford University Press Inc, 1970.

---. *The Language of Morals*. Oxford: Clarendon Press, 2003.

Mackie, J.L.. *Inventing Right and Wrong*. London: Penguin, 1990.

Plato. Edman, Irwin ed. *The Works of Plato*. New York: Tudor Publishing Company, 1936.

Tait, Katharine. *My Father, Bertrand Russell*. London: Thoemmes Continuum, 1995.

