



Free Irrationality in Moral Choices

Minji Jang
Carleton College

This paper explores the question of whether an agent can freely and willfully act against her all-things-considered best judgment. I construct the case of an agent acting against the demands of impartial morality that give her decisive reason to act, without compromising her full agential power and strong will. I call this her demonstration of free irrationality and distinguish it from the accounts of weakness of will and volitional necessity. The questions are whether there is a kind of non-rational motivational force strong enough to cause us to act irrationally, and whether an action motivated by such force is compatible with our exercise of full agential power. I offer one possible explanation: the strongest and deepest commitments central to our lives and personal characters may motivate us to act contrary to our rational judgments. Then the questions are whether it is possible for such strongest commitment not to give us decisive reason to act, and whether we can still act on that commitment, knowing that it does not give us decisive reason to act. I conclude that if such a case were possible, then it would show that our exercise of agency does not always solely depend on our rationality and that the former can be dissociated from the latter, at least on some rare and extreme occasions.

I. Case Study of Annie

Annie deeply loves her son. Her world crumbles when she finds that her son has committed a murder and the police are looking for him. If she tells the truth and turns him in, he will spend the rest of his life in jail, which will seriously compromise his welfare. If she lies and hides her son, the police will arrest the innocent man living next door instead.¹ Annie does recognize that in the given situation, impartial morality demands that she tell the truth; she is aware that it is morally right to turn him in so that he pays for the crime he has committed. If, nevertheless, she chooses to hide her son, how can we make sense of her behavior?

There are two easy explanations. In the first account, Annie is weak-willed at the moment and acts against her best judgment under compulsion; in this case, she does not exercise full agential power, which I define below. In the second account, Annie finds that morality does not give her decisive reason to act and judges that it is best to hide her son; in this case, she exercises full agential power and acts rationally in accordance with her best judgment. While

it might seem as if these two are the only conceivable explanations for Annie's behavior, this paper aims to examine the possibility of an alternative account.

In this third account, Annie judges that moral demands give her decisive reason to tell the truth and retains her full agential power in acting contrary to that judgment. Before I move on, I define *full agential power* and *all-things-considered best judgment* as follows:

D1: An agent exercises *full agential power* in ϕ -ing if, and only if, at the time she ϕ , she acts without compulsion or coercion and brings about ϕ as a result of her strong will.

D2: An agent has a *decisive reason* to ϕ if, and only if, she judges that all things considered, it is best to ϕ ; when forming this *all-things-considered best judgment*, she takes into consideration all she knows, believes, and holds at the time she ϕ , including all the relevant principles, beliefs and desires.²

I further define an agent's performing rational and irrational actions as follows:

D3: An agent *acts rationally* if, and only if, she acts in accordance with her all-things-considered best judgment; she *acts irrationally* by either not taking the action directed by her best judgment or by taking an alternative course of action.

In exercising full agential power, Annie is neither coerced by external force nor overcome by compulsion *from within her*, such as a surge of feelings or physical and mental illness. A drug addict, for instance, fails to exercise full agential power, for he is driven by his compulsive desire to take drugs that he knows to be detrimental to his health and overall happiness. As Harry Frankfurt explains, while the forces of addiction are still "generated within him," they are "nonetheless not in the fullest sense his own," in that those are the "forces with which he does not identify, whose influence he struggles to resist."³ The addict acts against his judgment that it is best to stop taking drugs, with his mental faculty and clarity compromised by the forces of addiction, and fails to bring about the action as a result of his strong will.⁴

In contrast, bringing it about as a result of her strong and free will, Annie gives *full consent* to her action; instead of struggling to resist its influence, she accepts and endorses the motivational force of the action she chooses to perform.⁵ The question is whether she can give this kind of consent to an action contrary to her best judgment. In my alternative account, Annie gives full consent to her decision to hide her son, despite her judging that all things considered, it is best to turn him in. By acting against her judgment, Annie acts irrationally; by maintaining full agential power, she acts freely. I call this her demonstration of *free irrationality*.

The goal of this paper is to examine the possibility of this account. The second and third sections introduce the traditional explanations of an agent acting against moral demands, the cases of *weakness of will* and *volitional necessity*. The fourth and fifth sections attempt to construct a distinct account of *free irrationality* and answer the potential objections. The last section concludes with what the possibility of this account reveals about the nature of our agential power and rationality. It should be acknowledged that the case study

discussed in this paper demonstrates a specific case of conflict between the demands of love and morality. The explanations may bring in complications based on the nature of these demands, which apply exclusively to this case and not to more general cases. If we successfully construct the case of Annie, however, we can show that an agent performing a *freely irrational* action is at least possible in some cases of practical conflict and opens up an investigation in a broader context.

II. Weakness of Will and Acting Under Compulsion

The first account explains that in acting contrary to moral demands, Annie shows weakness of will. Annie judges that it is best, all things considered, to tell the truth, yet she is weak-willed at the moment and yields to the temptation to lie. Donald Davidson defines that a person acts incontinently if, and only if, he intentionally takes a certain action x , while judging that there is an alternative action y that he believes is open to him, and that, all things considered, it would be better to do y than to do x .⁶ Let me introduce Davidson's example of an agent acting incontinently and call him Bob. Only after lying on his bed does Bob realize that he has not brushed his teeth, and now he faces a conflict between brushing his teeth and staying in bed.⁷ Suppose that Bob judges that, everything considered, he ought to brush his teeth; his concern for health and the potential cost of dental treatment gives him decisive reason to rise from his bed.⁸ Nonetheless, he ends up staying in bed. By acting intentionally against his best judgment, he acts incontinently; by surrendering to temptation and acting in accordance with immediate pleasure rather than reason, he demonstrates weakness of will.⁹

This first account claims that this is precisely what happens to Annie. Perhaps she is under extreme stress from the police interrogation or is overcome with fear and anxiety about the possibility of her son spending devastating time in jail. While, based on all her principles, beliefs, and desires at the moment, she judges that she ought to turn him in, Annie ends up acting contrary to her judgment.¹⁰ It is crucial to notice that when Annie forms her all-things-considered best judgment, she does incorporate the desire to promote her son's welfare. She reasons that, all things considered, it is still better to take an action compromising that desire than to take an action promoting it. By acting to promote that desire, then, she ends up taking what she knows to be a worse course of action, thereby acting irrationally.¹¹ If asked, Annie would readily admit that she gave in to temptation or that she made a rash decision in the heat of the moment.¹² She might even admit that she deeply regretted her action of lying and sending the innocent man to jail and that she would not act in the same way if the same situation arose in future.¹³

Defining agency as the capacity to act in accordance with one's best judgment, the traditional view insists that the fact that Annie has acted against her best judgment reveals that she was weak-willed at the moment.¹⁴ Of course, this is not the exclusive account. Another easily conceivable explanation for Annie's choosing to hide her son is to claim that she was not, in fact,

acting against her best judgment; Annie might have judged that, all things considered, it is best to abandon the demands of morality and instead follow the demands of love. The next section examines this second account, in which Annie maintains full agential power and strong will in acting against moral demands by rationally judging that it is the best thing to do.

III. *Volitional Necessity* and Having Decisive Reason to Act Against Moral Demands

In the second account, Annie acts in accordance with her best judgment by acting contrary to what impartial morality demands. She still acknowledges that it is morally right to tell the truth and takes such a principle into consideration in forming her best judgment, yet she judges that it is not the demands of morality but the demands of her love for her child that give her the strongest, most decisive reason to act. By taking into consideration all her principles, beliefs, and desires, Annie judges that it is best to do what is in the best interest of her child and that is to hide him from the police.¹⁵ Susan Wolf explains that this is what happens in her original radical choice example, in which Annie decides that moral considerations “pale in significance” compared to her child’s welfare. Annie exercises full agential power, for she acts without compulsion or coercion and brings about the described action as a result of her strong will¹⁶. If asked, she would confidently say that she did not regret her action and that she would repeat the same action if the same situation arose in future, swearing, “I will go to hell if I have to, but my son is more important to me than my moral salvation.”¹⁷

The key idea behind this explanation is that Annie’s love for her child gives her, from her perspective, overriding reasons to act in favor of his welfare, in which case she acts rationally against moral demands. Harry Frankfurt takes a similar approach in his discussion of *volitional necessity*, in which he claims that, when people deeply and sincerely care about someone, this love gives people overriding reasons to act in accordance with the interests of the loved ones.¹⁸ He explains that “an encounter with necessity of this sort . . . somehow [makes] it apparent to [the agent] that every apparent alternative to that course is unthinkable.”¹⁹ If Annie were under the force of *volitional necessity*, then she would acknowledge that her love for her son gives her “powerful and often decisively preemptive reason for performing [the] action” that benefits him, which, in this case, is to hide him from the police so that he avoids spending a lifetime in jail.²⁰

How is the force of *volitional necessity* different from the force of compulsion, such as addiction, anxiety, and mental illness? According to Frankfurt, the key difference between an agent acting under the constraints of *volitional necessity* and a weak-willed agent acting under compulsion is that in the former case, the agent accedes to the force, not “because he lacks sufficient strength of will to defeat it,” but “because he is unwilling to oppose it” and “is unwilling to alter” his unwillingness; in this sense, the force of necessity is “to a certain extent self-imposed” by the fully empowered agent.²¹ Indeed, it is the

agent's own endorsement of the force that makes it irresistible and makes every alternative course of action unthinkable to him, which is distinct from a case of an agent unable to act otherwise due to the force of addiction or other types of compulsion, which he struggles to resist.²²

The compatibility between such constraints and full agential power will continue to be discussed in the following sections. For now, I summarize the two traditional accounts of Annie's acting contrary to moral demands. In the first account, she shows weakness of will, through which she fails to exercise full agential power and acts compulsively. In the second account, she acts with her full agential power, for she judges that the demands of love give her decisive reason to act against impartial moral demands. In regard to this second account, she may have been under the constraints of *volitional necessity*, so that her deep caring for her son renders all alternative courses of action, potentially compromising his interests, unthinkable. The question is whether there can be an alternative account of Annie's behavior, distinct from these two explanations. More specifically, is it possible for Annie both to judge that it is best to turn her son in and retain full agential power in acting contrary to that judgment?

IV. Attempts at Free Irrationality

The first step towards constructing the case of *free irrationality* is to distinguish it from the previously described accounts. Again, by performing a *freely irrational* action, an agent voluntarily and willfully acts against her best judgment, formed on the basis of her deliberating about all relevant principles, beliefs, attitudes, and desires she has at the moment. Furthermore, she exercises full agential power, through which she acts without compulsion or coercion, but with her strong will to perform that action. She does not make a decision to hide her son in the heat of the moment or as a result of being compelled by extreme fear and anxiety. She still feels fear and anxiety at the prospect of her son's imprisonment, but such an emotional state is neither the exclusive nor the strongest motivational force of her action.

Instead of surrendering to temptation, she acknowledges the kind of pleasure to be gained from lying, takes the desire to have that pleasure into consideration, and judges that hiding her son would still be a worse course of action, all things considered, than turning him in. She may believe and reasonably judge that the demands of impartial morality always give her decisive reason to act. If, nevertheless, she voluntarily and willfully chooses to hide her son, then this case calls for a separate account of her behavior, in which she retains full agential power in acting contrary to her best judgment, thereby acting *freely irrationally*.

How can this be possible? I direct our attention to what can possibly motivate an agent to act contrary to her rational judgment.²³ Some motivational force causes the agent to abandon her reason, yet she chooses not to give up the source of such motivational force. While I do not claim this to be the exclusive or exhaustive source, I propose one possibility that such motivational force is

generated by an agent's strong commitment to someone about whom she cares deeply. In this regard, an agent performing a *freely irrational* action, in the described conflict between the demands of love and morality, is similar to an agent acting under the force of *volitional necessity*.²⁴ Both are strongly, even irresistibly, motivated by their love for the loved ones.

This idea of agents voluntarily and willfully imposing powerful constraints on themselves gives us a valuable insight into the case of agents performing a *freely irrational* action. The crucial difference between the two accounts is that agents under the force of *volitional necessity* believe and admit that their love gives them decisive reason to act in favor of their loved ones. By taking the actions that they judge to be in the best interest of their loved ones or by refusing to take the actions that they judge to compromise the loved ones' interests, the agents act rationally. While Annie's action can also be driven by her strong commitment to her son whom she deeply loves and about whom she cares, she does not find that this commitment gives her decisive reason to act. Her acting on this commitment is *freely irrational*, for she knows that such commitment fails to give her decisive reason to act, but she still freely acts on it.

I further suppose that not every commitment can generate strong enough motivational force to cause an agent to act irrationally. Perhaps only the strongest and the deepest commitments are capable of this power, and the agent chooses not to give them up in any circumstance, despite the possibility of her being motivated to act irrationally. The questions are whether it is possible for the strongest form of commitment not to give the agent decisive reason to act, and whether the agent can still act on that commitment, knowing that it does not give her decisive reason to act. The following section attempts to address and answer these questions.

V. Questions and Answers

The assumption is that there are strong and deep commitments that are central to Annie's life, such as her commitment to loving and protecting her son. This kind of commitment renders her unable to perform an action she knows to violate the commitment, such as one that severely compromises her child's welfare. Annie does not make an impulsive decision in the heat of the moment, nor is she coerced by some external force. She knows that her commitment to care for her son will potentially render other courses of action, including what she reasonably judges to be the best, unthinkable. Again, the two main questions are whether this kind of strong commitment can fail to give Annie decisive reasons to act, and whether she can still choose to act on that commitment without losing her agential power.

The traditional accounts seem to suggest that an agent's acting against her rational judgment and exercising her full agential power are mutually exclusive, in that she necessarily has to sacrifice her agency, at least to some extent, when acting irrationally. If, however, the agent is capable of *both* acting against her reasoned judgment *and* maintaining full agential power, then it

would suggest that a *freely irrational* action is possible and that, at least on some rare occasions, her exercise of agency can be separated from her dependence on rationality.

Returning to the two questions at hand, I begin with the question of how the agent can choose to act on the commitment, if she knows that it does not give her decisive reason to act. I first clarify what I mean by an agent acting on her commitment:

D4: If an agent is *committed* to ϕ , then she will robustly ϕ in a particular situation S, for part of her being *committed* to ϕ is robustly ϕ ing in S.

In robustly performing an action, an agent endorses her action and simultaneously finds other courses of action unthinkable to her. Furthermore, in endorsing her action in this way, Annie not only desires to perform the described action, but also endorses her desire to perform the action. Borrowing Frankfurt's notions of first-order and second-order desires, Annie has both first-order desire to ϕ and second-order desire to desire her ϕ -ing, the former simply being her desire to act or not to act and the latter being her desire to have certain desires and motivations.²⁵

An agent's acting robustly in this sense, however, does not require that she is *fully integrated* or *wholehearted* in performing the chosen action.²⁶ She may not have resolved all conflicts among her second-order desires by identifying herself with a single desire that constitutes her will. In contrast, in robustly choosing to hide her son in the given situation, Annie still has conflicts among her second-order desires, in that she not only desires to be motivated by her desire to protect her son's welfare, but also desires to desire to act morally or to not act immorally.²⁷ She does not have a similar kind of conflict among her first-order desires, however, for she finds performing the alternative course of action, such as turning in her son, unthinkable; she has no desire to turn him in, yet she has a desire to desire to turn him in. It is only among the higher-order desires that she experiences an absence of wholeheartedness, and this does not deter her from robustly performing the action with full agential power.²⁸

Most of times, if not always, Annie would not encounter this type of conflict. There need not be systematic conflicts between the demands of impartial morality and the demands of love, nor should there be necessary or frequent conflicts between her commitments and rationality. Instead, what her commitment demands and what her reason demands would normally be aligned with each other in a way that she can rationally act to protect the welfare of her son. While it is possible that Annie would never have faced this type of tragic scenario where her son becomes a murderer and lays his fate in her hands, Annie is put in the much-less-ideal situation and now has to make a radical choice between following her rational judgment and acting on her commitment. The question is whether she can choose the latter at the expense of the former without compromising her agency.

Furthermore, unlike the case of a *wholehearted* agent, Annie's following her commitment and robustly hiding her son in this particular circumstance

does not entail her performing the same action in any and all situations where her son's welfare is at risk. If Annie faces a situation in which she has to choose between saving her son and exterminating the human race, then her commitment to loving her son may not entail her robustly choosing to protect his welfare. I propose to conceive of a kind of commitment that she has toward loving her son, so that, when she has to choose between turning in her son and turning in the innocent man living next door, she robustly chooses the latter action against her rational judgment to take the former.²⁹

Why does Annie not just abandon her commitment or the way she is committed? I propose one possible explanation: she chooses not to give up the kind and degree of commitment she has, despite the possibility of being motivated to abandon her best judgment on some occasions, perhaps because she knows that it is the strongest and deepest commitment she has toward the person she loves most in this world and that keeping such commitment *the way it is* forms an indispensable part of herself.³⁰ This is not to say that Annie's action is an ideal way of loving her child, or that every parent deeply loving his or her child is necessarily committed this way. It is the way Annie loves her child, and she cannot both love him the same way that she does now and abandon the kind and degree of commitment she has toward protecting him.

She cannot abandon this commitment, again perhaps because this particular way of loving her son is so central to her character. It is important to understand that she is not acting *in order to* preserve her character; it is not her primary goal. Rather, the action that she chooses to perform is an expression of her character, the kind of person she is and perceives herself to be, which she accepts and with which she lives. Loving her son the way she loves him now forms an integral part of her character, and such commitment is inseparable from her robustly protecting his welfare in the given situation. Annie would say, "I know that I am acting irrationally, but *I live with it*. I cannot love my son any other way than I love him now, and I cannot send him to jail while loving him the way I do." In this way, she acts irrationally against her best judgment, yet she does so freely and willfully. This possibility suggests that Annie need not sacrifice her agential power in acting contrary to her rational judgment, which implies that her agential power is not necessarily, hence not always, closely tied to her rationality.

I believe that Christine Korsgaard's discussion of practical identity would help us understand how an action constitutes an integral part of an agent's character, and vice versa.³¹ Korsgaard claims that what we deliberately choose to do decides who we are, and that in choosing to perform those actions, we create our practical identity. We constitute ourselves as the authors of our actions by the very act of choosing them³², and such actions springing from our own personal characters and identities imply our exercise of uncompromised agency, despite the actions not being directly filtered through our reasons. In acting contrary to her rational judgment, Annie is still "the possessor of [her] personal or practical identity" and "the author of [her] action," and this is again how her action is distinguished from the case of compulsion.³³

Going back to the second question, if such commitments are so central to Annie's life, how is it possible that they still fail to give her decisive reason to act? In many cases, the agent's deepest commitments give her decisive reason to act; the possibility of *free irrationality* denies that it is necessarily true in all cases. It allows that the agent still robustly ϕ in S , even when she knows that *before* S arises and *during* S , ϕ -ing is not what she judges to be the best, all things considered. Annie may not even know, until the very moment she faces the situation, that she is committed to loving her son in a way that she would robustly hide him from the police in the given situation. Or, she may already know that her commitment will, in rare and extreme situations like this one, demand that she act irrationally. She truly hopes that such situations do not arise, so that she need not make a choice between following her reason and abandoning her commitment. Yet, when the situation does arise, she chooses not to give up the commitment and robustly acts on it.

Still, how is it possible that Annie rationally judges that it is best, all things considered, to take an action that she knows to sacrifice her son's welfare, when she loves him so dearly? Suppose Annie has always believed that it is not best, all things considered, or even certainly worse to acquit anyone of a crime, including the person whom she loves the most in the world. I propose that it is conceivable and possible that not only *before*, but also *during* the time she makes a decision to hide her son, Annie *sustains her judgment* that what she does now is not the best thing to do, all things considered. If she sustains this judgment, then her commitment fails to give her decisive reason or shift her judgment of what gives her decisive reason to act.

In performing a *freely irrational* action, an agent exercises full agential power and sustains her judgment about the values of conflicting options. Annie sustains her judgment that she now takes a worse course of action, all things considered. She admits the presence of dissonance that her desire not to abandon her commitment is at odds with her rational judgment. She wishes to avoid this kind of dissonance and not to face conflicts causing such dissonance, yet, when they do arise, she neither succumbs to temptation nor abandons her commitment; she faces the inevitable dissonance and freely acts against her rationality.

VI. Conclusion

In this paper, I proposed the possibility of *free irrationality* exercised by an agent facing a conflict between the demands of love and impartial morality. I attempted to construct a case of an agent freely acting against her all-things-considered best judgment, incorporating all accessible principles, beliefs, opinions, and desires. I claimed that she exercises full agential power in acting contrary to her judgment in that she brings about her action as a result of her strong will. I distinguished this account from the standard case of weakness of will, in which an agent acts contrary to reason under compulsion, and also from the original case introduced by Susan Wolf, in which an agent finds over-

riding reason to follow the demands of love, hence acting rationally in abandoning moral demands.

In constructing the account of *free irrationality*, I offered one possible explanation: Annie's deep commitment to loving her son motivates her to act against her best judgment. She is committed in such a way that she robustly takes an action to protect him in the given situation, despite the fact that such commitment does not give her decisive reason to act. She cannot love her son the way she does and abandon this commitment, yet loving him the way she does is so integral to her life and personal character. I further suggested that if Annie were to sustain her judgment that she now takes a worse course of action at the moment of performing that action, then it would show that even her strongest commitment fails to give her decisive reason to act or shift her judgment of what is the best thing to do, all things considered. By acting robustly on the commitment that fails to give her decisive reason to act, Annie acts *freely irrationally*.

It is important to note that the possibility of *free irrationality* commits us to the view that our exercise of full agential power can be divorced from our dependence on rationality on some extreme occasions. In other words, we can bring about the action as a result of our strong will but not exclusively through our reason; we can be fully empowered agents and freely act against what our rationality directs us to do. This suggested possibility of *free irrationality* and its implications, however, does not raise great concerns for our performance of rationality. As David Pears points out, this kind of action is irrational, *only* in the sense that it is "not an element in the agent's reasoning" or is "ruled by [the] edict of reason."³⁴ The possibility that there are values and commitments, central to our lives and characters, which may motivate us to act without necessarily giving us decisive reason to do so, does not compromise our status as rational agents.

Notes

- 1.) This case study is constructed based on the radical choice example introduced by Susan Wolf, "Morality and Partiality," *Philosophical Perspectives* 6, Ethics (1992): 243-259. Annie's choice has to be radical, for she faces a conflict between the demands of impartial morality and the demands of love, and, by acting in accordance with one, she necessarily has to abandon the other. While some, especially the guardians of Kantian perfect duties, may deem a simple act of lying immoral, Wolf adds an additional consequence to Annie's action to highlight its perceived immorality. If she lies to the police in order to save her son, she now sends an innocent man to jail for the rest of his life. Sending an innocent man to jail for the rest of his life is considered acting against the demands of impartial morality; sending her son to jail for the rest of his life is considered acting against the demands of love. Hence, Annie has to make a radical choice between the two demands, both of which she values.
- 2.) Donald Davidson, "How is Weakness of the Will Possible? (1969)," in *Essays on Actions and Events* (Oxford: Clarendon Press, 2001), 21-42. See "the phrase 'all things considered' must, of course, refer only to things known, believed, or held by the agent, the sum of his relevant principles, opinions, attitudes, and desires" (40). I also assume that when an agent forms her all-things-considered judgment with full agency, she is not being willfully ignorant and neglecting information or facts relevant to her decision. Annie would not neglect the fact that by lying to the police, she would send an innocent man to jail, nor would she ignore that doing so would be counter to moral demands.
- 3.) Harry G. Frankfurt, *The Importance of What We Care About*, 15th ed. (New York: Cambridge University Press, 2009), 183.
- 4.) Frankfurt does acknowledge and discuss the case of a willing addict, who not only endorses his addiction, but also would do whatever he could to maintain the force of his addiction if it were to fade. Yet, Frankfurt explains that "the willing addict's will is not free, for his desire to take the drug will be effective regardless of whether or not he wants this desire to constitute his will" (Ibid., 24). Although the willing addict chooses to take drugs, he does not bring about the action of taking drugs as a result of his strong and uncompromised will. It is not his endorsement that brings about the action, for he will take drugs regardless of his endorsement.
- 5.) Ibid., 182. This idea that an agent's endorsement distinguishes acting under volitional necessity from acting under compulsion will be further discussed in the later sections.
- 6.) Davidson, 22.
- 7.) Ibid., 30.
- 8.) In the original example, the agent judges that he would do better to stay in bed, everything considered, and then acts against that judgment; I adjusted the example so that it looks more similar to our case study.
- 9.) Ibid., 35.
- 10.) Whether Annie's judgment is indeed correct or whether the action against that judgment is indeed wrong is irrelevant to our discussion of *free irrationality*.
- 11.) I assume that it is possible for an agent to act against his or her best judgment, regardless of seeing it as a case of akrasia. See, for example, Richard Holton, "Intention and Weakness of Will," *Journal of Philosophy* 96, no. 5 (May, 1999): 241-262, and Richard Holton, "Inverse Akrasia and Weakness of Will," <http://web.mit.edu/holton/www/pubs/InverseAkrasia.pdf> (accessed March 18, 2015) for arguments against this assumption.

12.) Davidson gives several psychological explanations for an agent acting against his best judgment, such as self-deception, overpowering desires, and lack of imagination (42). He concludes that the agent has no reason for taking what he believes to be a worse course of action, which is the starting point of our discussion. What non-rational motivational force can be strong enough to cause the agent to act irrationally? Is such kind of action truly compatible with his exercise of full agential power?

13.) This is not to say that a weak-willed agent necessarily regrets his action. I point to the possibility of a weak-willed agent regretting the action that he performed under compulsion, only in the sense that he might believe that he would not have acted the same had he retained full strength of will. It is possible for an agent not to regret the action that he performed with weak will or to regret the action that he performed with full strength of will.

14.) Davidson seems to deny this view, for he acknowledges the possibility of a strong-willed irrational action. To the question of whether an agent can “have an unclouded, unwavering judgment that [his] action is not for the best, all things considered” and still perform the action with “no hint of compulsion or of the compulsive,” Davidson answers that “there is no proving [that] such actions exist; but it seems to me absolutely certain that they do” (29). The question is whether Davidson is right and whether one can at least explain, through a discussion of *free irrationality*, how such actions may be performed.

15.) I do not propose that excusing a child for the crime he has committed is indeed in his best interest; in this account, if Annie were to judge that it is the best way of promoting her son’s interest, then she would judge it to be the best thing to do, all things considered.

16.) Wolf, 254.

17.) *Ibid.*, 254.

17.) See, for example, Frankfurt, *Importance, and Taking Ourselves Seriously and Getting It Right* (Stanford: Stanford University Press, 2006).

18.) Frankfurt, *Importance*, 86.

19.) Frankfurt, *Taking Ourselves Seriously*, 42. Again, I do not propose that excusing a child for his crime is the way to benefit him the most. If Annie were to believe that it is the action that benefits him, then, under the force of volitional necessity, she would take that action and render all other courses of action unthinkable.

20.) Frankfurt, *Importance*, 87.

21.) *Ibid.*, 182. It is possible for an addict to endorse his or her addiction. (See footnote 4 for the description of a willing addict.) However, as Frankfurt points out, it is not his endorsement that makes the force of addiction irresistible; it is the addiction itself that is irresistible, regardless of whether the agent gives his consent to being constrained by it. This is what distinguishes the case of volitional necessity from the case of compulsion or addiction.

22.) It is still disputable whether anything beyond reason can motivate us to act without compromising our agency; traditional literature often argues that acting rationally is precisely what defines our agency, hence considering irrational actions as proof of our weakened agency. The goal is not to prove the counter-case empirically but to show how the counter-case would look if it were to happen.

23.) As I explained above, this source of motivational force may apply exclusively to the case of an agent performing a freely irrational action in the described conflict between the demands of love and morality. I open up the possibility of a freely irrational action performed by a fully empowered agent in other cases, although this paper aims to examine one specific type of conflict.

24.) Frankfurt, *Importance*, especially Chapter 12, "Identification and Wholeheartedness," 159-176.

26.) *Ibid.*, Chapter 12, "Identification and Wholeheartedness," 159-176.

25.) It is possible that Annie has not resolved conflicts among her higher-order volitions and desires, yet is it a necessary condition for her performing a freely irrational action? In other words, is it necessary that Annie have some incoherence, inconsistency, or conflicts among her higher-order preferences for her to act against her best judgment and acknowledge that it is irrational? I am inclined to say yes to this question, for if she were fully integrated and eliminated all conflicts within her, then she would be likely to find her action in accordance with her rationality. Perhaps it is because she does not wholeheartedly identify herself with a single desire that she finds that such desire does not give her decisive reason to act.

28.) Frankfurt may disagree with this statement, for he believes that an agent cannot have second-order volition if there remains an unresolved conflict among his second-order desires (*Ibid.*, 21), and that having a second-order volition is essential to her being a person (*Ibid.*, 16). If what distinguishes second-order volitions from the rest of second-order desires is that the former are "[desire to have] a certain desire to be his will," while the latter are simply "[desire] to have a certain desire" (*Ibid.*, 16), then I claim that a formation of second-order volition of this kind does not require the absence of all conflicts among second-order desires. An agent can desire to have a certain desire that constitutes her will, although not exclusively and wholeheartedly, thereby forming her higher-order volitions.

26.) This is not to say that sending one innocent man to jail is any less immoral than exterminating the human race in an objective sense. I only point out that Annie's commitment to protecting her son need not and probably would not entail her acting for his welfare in any and all situations. Instead, such commitment would demand that she act on his behalf in particular situations like the one she now faces.

30.) This idea springs from Michael Slote's discussion of one's action that cannot be conceptually prised from one's trait. See, for example, Michael Slote, "Admirable Immorality," in *Goods and Virtues* (Oxford: Clarendon Press, 1983), 77-107.

31.) Christine M. Korsgaard, "Agency and Practical Identity," in *Self-constitution: Agency, Identity and Integrity* (Oxford: Oxford University Press, 2009), 18-26.

32.) *Ibid.*, 20

33.) *Ibid.*, 20. Our discussion of free irrationality departs from Korsgaard's, in that she believes that our practical identities are "the sources of our reasons," and that "morality itself is grounded in an essential form" of such identities (*Ibid.*, 21-22). The possibility of freely irrational actions denies that our practical identities, which shape and are shaped by the actions that we choose to perform, are always the sources of our reasons, for such possibility allows for the actions motivated by our personal identities yet are not necessarily filtered through our reasons. If such conflicts calling for freely irrational actions were to arise only on rare and extreme occasions, Korsgaard would be right in claiming that those identities are generally the sources of our reasons, although perhaps not always.

34.) David F. Pears, "Motivated Irrationality," *Proceedings of the Aristotelian Society, Supplementary Volumes* 56 (1982), 167. See specifically: "...the suggestion to be developed is that its irrationality is not like the irrationality of perverse belief-formation. That involves, in the limiting case, believing something impossible, but the irrationality of perverse action is quite different. The action is irrational because it is unreasonable and it is unreasonable only in the sense that it does not obey reason."

Bibliography

- Davidson, Donald. "How Is Weakness of the Will Possible? (1969)." In *Essays on Actions and Events*, 21-42. Oxford: Clarendon Press, 2001.
- Frankfurt, Harry G. *The Importance of What We Care About*. 15th ed. New York: Cambridge University Press, 2009.
- , *Taking Ourselves Seriously & Getting It Right*. Stanford: Stanford University Press, 2006.
- Holton, Richard. "Intention and Weakness of Will." *The Journal of Philosophy* 96, no. 5 (May 1999): 241-262.
- . "Inverse Akrasia and Weakness of Will." <http://web.mit.edu/holton/www/pubs/InverseAkrasia.pdf> (accessed March 18, 2015).
- Korsgaard, Christine M. "Agency and Practical Identity." In *Self-constitution: Agency, Identity and Integrity*, 18-26. Oxford: Oxford University Press, 2009.
- Pears, David F. "Motivated Irrationality." *Proceedings of the Aristotelian Society, Supplementary Volumes* 56 (1982): 157-177.
- Slote, Michael. "Admirable Immorality." In *Goods and Virtues*, 77-108. Oxford: Clarendon Press, 1983.
- Wolf, Susan. "Morality and Partiality." *Philosophical Perspectives* 6, Ethics (1992): 243-259.