ESSAYS ON OBLIGATION

by

Johanna Goth Finegan

A.B., Philosophy
Princeton University, 2000

SUBMITTED TO THE DEPARTMENT OF PHILOSOPHY IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY IN PHILOSOPHY AT THE MASSACHUSETTS INSTITUTE OF TECHNOLOGY SEPTEMBER 2007

©2007 Johanna Goth Finegan. All rights reserved.

The author hereby grants to MIT permission to reproduce and to distribute publicly paper and electronic copies of this thesis document in whole or in part in any medium now known or hereafter created.

Signature of Author: _______________________________ Department of Philosophy September 7, 2007

Certified by: _______________________________ Judith Jarvis Thomson Professor of Philosophy Thesis Supervisor

Accepted by: _______________________________ Alex Byrne Professor of Philosophy Chairman, Committee for Graduate Students
Essays on Obligation

by

Johanna Goth Finegan

Submitted to the Department of Linguistics and Philosophy on September 7, 2007 in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy in Philosophy

ABSTRACT

This three-chapter thesis explores philosophical issues in the area of ethical and epistemic obligations.

In the first chapter, “Supererogation and Defenses of Modest Moral Demands,” I argue that a range of recent approaches to defending the “ordinary morality” view that the demands of morality are fairly modest make it hard to see how there could be any supererogatory acts, and therefore sacrifice one aspect of our intuitive view of morality in an attempt to preserve another. I go on to consider some difficulties in accommodating supererogation that defenders of “ordinary morality” might face more generally.

In the second chapter, “Deontological Judgments about Belief,” I address the problem of how deontological judgments about belief can ever be true, given that the apparently involuntary nature of belief seems incompatible with the sort of responsibility needed for deontological judgments. I consider and reject arguments which purport to show that we have the right sort of voluntary control over beliefs, as well as arguments which purport to show that considerations of voluntary control are altogether irrelevant to whether deontological judgments about belief can be true. I argue instead that deontological judgments about belief can be true in virtue of our participation in self-deceptive practices, something over which we do have the right kind of control.

In the third chapter, “Belief and Commitment” I consider arguments to the effect that we must sometimes violate epistemic norms in order to enjoy the goods of friendship and other social relationships and in order to give those we are in relationship with their due. I raise some worries for such views and then argue that we can enjoy the goods and fulfill the obligations of social relationships without paying such a price, that an alternative attitude, which I call commitment, can get us at least most of the riches of social relationship without the costs of irrational, self-deceived belief.

Thesis Supervisor: Judith Jarvis Thomson
Title: Professor of Philosophy
CONTENTS

1. Supererogation and Defenses of Modest Moral Demands  4
2. Deontological Judgments about Belief             30
3. Belief and Commitment                           66
Supererogation and Defenses of Modest Moral Demands

1. Introduction

1.1 Ordinary morality and moral requirements

Let us take “ordinary morality” to be not a particular theory, but a loose and general characterization of the sorts of moral intuitions and judgments “people like us” tend to have and make. Ordinary morality (OM), at least at first glance, seems to make fairly modest claims on us in terms of what we are obligated to do to promote the welfare of others. We need to do something for other people, a bit toward making the world a better place, but not so very much that it would require most of us to lead lives radically different from our current ones. For example, OM might tell us that at some point, when we have the means to do so, we should contribute some of our financial resources to some charity or other, but it does not tell us that we ought to give up every cent left over after our minimal needs for sustenance are met to fight world hunger. OM might tell us that, in the absence of extremely pressing demands on our attention, if someone asks us for the time we ought to tell him, but it does not tell us that we have to give the questioner our watch. OM might tell us that we should spend a few hours per month volunteering in the service of some noble cause; it does not tell us that we ought to give up all other interests, projects, and relationships, to devote our entire lives to that cause.

The details of OM are somewhat hazy, as it is often controversial what the consensus is, if one exists at all. One might not agree with all of the above claims about what OM tells us, but I take it that they convey the general idea. Morality places demands on us, but these demands only go so far.
1.2 Challenges to Ordinary Morality

Consequentialism, at first glance, appears incompatible with the OM view that morality’s demands go only so far. Some consequentialists, like Frank Jackson, try to show otherwise by arguing that something like OM is actually what would have the best consequences. More often, though, consequentialists respond by challenging the view that the correct moral theory should fit with OM’s rather modest requirements, and argue that the demands of morality go far beyond the intuitive requirements. Following terminology used by one of the challengers, Shelly Kagan, I will call them extremists. In order to give a general idea of the sorts of challenges being posed, I will very briefly describe two representative examples.

Kagan in The Limits of Morality acknowledges that many of our intuitive ordinary moral judgments do suggest a morality with rather modest demands. But surely, he argues, we want more from a moral theory than a fit with our intuitions, as any unattractively ad hoc moral theory could do that. We want simplicity, coherence, and explanatory power. If the moral theories which provide us more of these desiderata require us to change some of our intuitive judgments, we should do so. Kagan takes considerable pains to show how most of the limit-explaining distinctions we use in trying to justify ordinary morality fail to stand up under closer scrutiny, and argues that there is no convincing principled way to draw the lines between obligation and permissibility where we want them.

Peter Unger’s approach in Living High and Letting Die is slightly different. Instead of arguing that the OM view of what morality requires of us is indefensible, he argues that the intuitions behind it are simply not moral at all, but are better ascribed to psychological factors.

---

2 Ibid., 11-14.
which distort our ability to respond to what he calls the Basic Moral Values. Some view these factors as shedding light on the content of morality, but Unger sees them as "noise" which should be disregarded. He offers explanations for why we tend to make the sorts of moral judgments that we do, and argues that those explanations do not make a moral difference.

The way to reply to these challenges is to articulate a principled, plausibly moral, justification for why the demands of moral obligation are roughly as we intuitively think they are. Unger offers non-moral explanations for our intuitions about cases; someone who wants to justify a modest view of moral requirements should try to offer moral ones. Kagan argues that there is no sound principled basis for making the distinctions we intuitively want to; someone who wants to defend the OM view should try to find such a basis.

Various theories can be seen as attempts to offer such a justification for a modestly demanding morality. Some are responding to explicit formulations of extremism, others responding to the potential extremism they see lurking in the foundations of utilitarianism. All are concerned to show that it is not the case that we ought to pursue a life of drastic self-sacrifice.

1.3 Supererogation

Another feature of OM related to modest moral demands is the concept of supererogation. A supererogatory act is an act which, while not obligatory, is permissible and good—in fact, very good. Our ordinary moral talk is full of dramatic examples of apparent supererogation, from the bystander who rushes into a burning building to save some children, to a rich person giving up his wealth to live humbly and serve the poor—as well as many less.

3 For a survey of many of these distorting factors, see Ch. 2 of Peter Unger's Living High and Letting Die: Our Illusion of Innocence (Oxford: Oxford University Press, 1996).
dramatic ones, including all who are a little more generous than they are obligated to be, as illustrated in the examples given in the introductory paragraph of this chapter.

OM does not merely allow for supererogation; traditionally, supererogation has been an integral part of our intuitive conception of morality. It seems that part of the reason why the OM defender would want to defend modest moral requirements would be to preserve the status of the supererogatory—to prevent the "moral saints" from turning out to be only slightly less miserable failures than the rest of us. One might argue that a view which supports OM demands has an advantage over an extremist view, precisely because it is not obviously incompatible with supererogation. Failure to accommodate supererogation just might be one more way in which extremism is unappealing to those who think that a moral theory ought to capture important intuitions about morality.

I will argue in this chapter that several defenses of OM (in particular those by Frank Jackson, Liam Murphy, Bernard Williams, and Susan Wolf) in favor of modest moral demands are no more compatible with supererogation than the extremist alternatives. Even if they are plausible justifications of modest moral demands, the distinctions they offer fail to account for supererogatory acts, and therefore they are unable (just like the extremists) to accommodate some of our basic intuitions about morality. Once such replies no longer have the advantage over extremism of fitting nicely with our ordinary moral intuitions, their appeal suffers considerable damage. In their attempts to show that we do not have to live an extremely demanding moral life, they make it very difficult to see how such a life could be good. A consideration of their arguments suggests that we must declare acts which go beyond what is morally obligatory to be defective in some respect, in order to defend our decision not to do them ourselves.
2. Getting clear about supererogation

2.1. The basics

Supererogatory acts are morally good acts (in fact, usually very morally good acts), which are permissible but not obligatory. It is not the case that we ought to do them, and it is not the case that we ought not do them. They involve going above and beyond the call of duty, doing more than is required, morally speaking. One deserves praise for doing the supererogatory—abundant praise, in fact; but does not deserve blame for failing to do it.

Attention needs to be given to the fact that morality is not silent about the value of these acts. Any account of moral requirements which views acts “going beyond” those requirements as being merely permissible has not succeeded in accounting for supererogation. It is not merely “okay” to live on half your income and donate the rest to famine relief—there is something morally excellent about doing such things.

While supererogatory acts surely deserve praise, they must be distinguished from those which are merely praiseworthy, admirable, or difficult, as those might well be obligatory. Sometimes the demands of morality, even on an OM view, are very demanding indeed. Depending on one’s role and circumstances, one may be obligated to do something which is extraordinarily hard for ordinary human beings to do. Extraordinary courage may be called for, for example, or a willingness to endure pain or death. OM does not deny that sometimes doing what we ought is very difficult and demanding; it only denies that it could be extremely demanding for all of us all the time.
2.2 Supererogation and generosity

For the sake of simplicity, my discussion in this chapter will focus on one kind of supererogation—that relating to beneficence (rather than, say, courage). I will focus particularly on cases involving giving of one’s resources.

Supererogatorily beneficent acts need to be distinguished from the merely generous ones. It might seem appealing at first to draw the dividing line between obligatory and supererogatory beneficence so that it coincides with the dividing line between generosity and justice (giving others what they have a right to). But the two do not match up. Especially where it does not cost us much to help, there are many kinds of assistance to others which it would be wrong for us to withhold from them, even though they do not have a right to our aid. If someone asks me the time, and my watch is right there on my wrist, it would be morally defective in me to refuse. If an acquaintance asks me if she can borrow an inexpensive, unremarkable book, and I know her to be trustworthy (we can suppose she has borrowed my friends’ books before, and has always returned them on time and in perfect condition), and I am not using the book myself, I do not do something supererogatory in lending her the book. I do not owe it to her, but she and others would be justified in thinking less of me if I refused without good reason.

Some acts of generosity are morally obligatory—even though no one has a right to them, we ought to do them, and are liable to moral criticism and blame if we do not, criticism for being what Thomson calls miserly. Another way of putting the point about generosity and supererogation is to say that the supererogatory acts of beneficence are those which are generous if we do them, but not miserly if we fail to do them.

2.3 Self-sacrifice

Intuitively we regard many acts of self-sacrifice, if at all successful, as good acts. Even if unsuccessful, we tend to give strong moral praise to the unfortunate self-sacrificers.

But we should note that not all self-sacrificial acts will be supererogatory. As I already mentioned, some might be morally obligatory. More strikingly, some might not even be morally good at all. They might involve insufficient attention to or value placed on oneself, one’s life, and one’s projects. We usually do not take this defect very seriously, if it is combined with altruistic motives. (It is something philosophers talk about when criticizing utilitarianism, but one rarely hears it as part of our ordinary practices of moral approval and disapproval.) We are more likely to think the act unfortunate or stupid than to think it morally wrong. But there may be something to be said for viewing such acts as morally defective.

There are two kinds of ways in which might we think self-sacrificial acts might not be good. The first kind involves cases where the cost to the agent seems unreasonably high compared to the benefit to the beneficiary of the act. People who sacrifice immensely in order to help others in small ways are not considered moral saints, they are considered foolish, at least when strangers are involved. The second kind involves cases where there would be a great benefit to the beneficiary of the act if it were successful, but the risk is extremely high and the chances of success are extremely low. My intuitions are somewhat mixed here, but it is at least plausible that someone who risks his life to save another’s with only a very tiny chance of success has done something that is not good. Therefore, I will not be considering these sorts of acts in my discussion of the supererogatory.
3. Accounts of Modest Moral Demands

3.1 Murphy

Liam Murphy’s argument for modest OM demands in the area of beneficence centers on his view of beneficence as a cooperative enterprise—we all share a collective duty to seek to improve the welfare of others.\(^5\) Extremist versions of what beneficence requires of us assume that promoting the welfare of others is an individual project, but it is more plausible to see us all as having the aim “to promote the good together with others.”\(^6\) Few if any of us actually have the individual duty. Extremist views of beneficence are thus too demanding because they expect us to take on the mission of the world’s welfare as an individual burden, regardless of what others do. This view allows Murphy to argue for a conception of one’s “fair share” of doing good—roughly, all the good that needs to be done, divided up amongst all the individuals in the group, according to their means. In trying to figure out how much we morally ought to give, we should consider how much we would need to give if everyone else were giving as well. Murphy argues that it is not the case that one morally ought to do more than this fair share, even if the others do not actually give anything. Their abdication of their responsibility does not add to yours.\(^7\)

In this way, Murphy appears able to get us modest requirements similar to those that ordinary morality calls for. It is plausible that if we all did our fair share, we could lead relatively normal lives, and only in very unusual circumstances would we have to make drastic sacrifices. Thus, the level of what we ought to do probably would not be too far off from what

---


\(^6\) Ibid., 286.

\(^7\) Ibid., 280.
we think it is. Of course, this is only a contingent matter. If there were only a tiny cluster of well-off people surrounded by millions of starving folk, each of them would be morally obligated to make drastic sacrifices in their lives in order to attempt to meet the needs of the others. OM would make demands on them far less modest than the demands it makes on us. This seems like the intuitively right result, especially if Murphy is correct that a main reason for our balking at the thought of having to make huge sacrifices is that we are surrounded by so many people not doing anything. While it is difficult to know what OM intuitions would say about such counterfactual situations, at the very least it seems right that OM demands would be responsive to the amount of need.

3.2 Williams

Bernard Williams does not address the issue directly, but his discussion of utilitarianism (in his “A Critique of Utilitarianism”) as it relates to integrity gives some hints of a possible approach one could take to justifying modest moral demands. He discusses the role that projects and relationships (valued and pursued for their own sake, not for moral reasons) play in giving meaning to human life. He argues that non-consequentialist commitment to projects and relationships give people’s lives what he calls “integrity”, something essential to a life worth living, and offers an argument against utilitarianism that works roughly like this: Promoting utility is a second-order project which requires first order projects to operate on. We increase utility by promoting people’s projects and relationships; but our projects and relationships would not have value if they were held for utilitarian motives. In this way, utilitarianism, or at least a

---

8 Ibid., 277-8.
10 Ibid., 110.
simple version of it, seems self-undermining. To accept utilitarianism as a guide for action is to thereby strip one’s projects and relationships of the value which utilitarianism aims to promote.

I will not address this argument further as it is not of much importance for the rest of the chapter. What is of interest is the role of freely chosen (non-morally motivated) projects and relationships—as these suggest a limit to moral demands, namely that taking on projects and relationships for non-moral reasons is essential for giving meaning and value to life. We had better see to it, on this view, that the demands of morality are restricted so as to leave room for these projects and relationships.

There are weaker and stronger versions of this claim. The weaker version says that in order to have integrity and a meaningful, happy life, one needs to have commitments which are not strictly based on maximizing utility—we need to be able to be committed to projects and relationships for their own sake to some degree, not because merely of their utility. The stronger version of the claim says that in order to have integrity and a meaningful, happy life, one’s commitments must be outside the realm of moral evaluation, utilitarian or otherwise. Real commitments, on this view, are commitments-whatever-the-consequences, and commitments-whatever-the-moral-implications. Peter Railton does a good job of arguing against this stronger version. We can, on his view, morally assess our commitments (and revise them accordingly) without draining them of all value.\footnote{Peter Railton, “Alienation, Consequentialism, and the Demands of Morality,” \textit{Philosophy and Public Affairs} 13 (1984), 150-2.} This seems right, so I will use the weaker interpretation throughout the rest of this chapter.

In sum, Williams suggests a justification for modest moral demands in his arguments for the importance of integrity. The demands of morality cannot be such as to threaten my giving proper independent value and weight to my projects and relationships.
3.3 Wolf

In “Moral Saints”, Susan Wolf argues that we should not admire or seek to become moral saints. Moral saints, according to Wolf, do not pursue non-moral interests and projects, or cultivate non-moral tastes. They do not develop an interest in literature, or a sport, or the arts. They are not able to enjoy gourmet cooking, or anything that cannot be justified “against morally beneficial alternatives”. In their aim to be perfectly kind, they end up being “dull-witted or humorless or bland”. Moral saints would not make good friends. Moral sainthood is disturbing because “it is an ideal of life in which morality unduly dominates”. In short, extremely moral lives are less valuable, less attractive human lives.

She has been criticized for offering an inaccurate description of what morally exemplary people are actually like and what moral perfection would look like. But I would suggest that her account is more or less fair in describing the maximally generous and beneficent, and some of her concerns really do seem to count against moral extremism. She may not have a case against moral theories in general, as she thinks, but she does point out some difficulties with the sort of life that appears to be recommended by the moral theories of people like Kagan and Unger. Perhaps the moral saint would not need to be as “bland” or “humorless” as Wolf suggests, but it is hard to see how relentlessly extreme concern for and devotion to improving the welfare of others could be compatible with the kind of investment in other interests and projects and relationships which make human lives attractive to us, the lives of the sort of people we want to be like and want to be friends with.

---

13 Ibid., 422.
14 Ibid.
15 Ibid., 424.
I should note that Wolf would object to being lumped in with the defenders of modest moral demands. For she does not take herself to show that true moral theories should only be so demanding. She is happy with the conclusion that morality is extremely demanding, but suggests that after a certain point we simply ought not follow its demands, that moral considerations are just one part of human life, which must not be overemphasized to the exclusion of other interests and practices. Morality itself might be very demanding, but one should only be so moral and not more.

Yet despite her protests, it looks as though Wolf puts forward a moral view of her own, if morality is understood in the broader sense of answering the general question of how we ought to live. To claim that a certain sort of life should not be admired or imitated is, in this broader sense, to make a negative moral judgment of it. We ought not, on her view, try to be "moral saints," or try to subject every area of our lives to moral scrutiny, on some specific views of what morality requires (she considers utilitarian and Kantian approaches). But is this not just to say that she thinks those approaches offer a deficient account of morality (i.e., they provide the wrong answers to the questions of what we ought to do and how we ought to live), and points toward another one instead?

On this interpretation, Wolf’s account does suggest a limit to the requirements of morality. Morality’s demands are limited where they interfere with our pursuit of the good life, or our being sufficiently interesting and personable to have our company enjoyed by others.

3.4. Jackson

Jackson’s “decision-theoretic consequentialism” also indirectly suggests a way of defending modest requirements that correspond roughly to those found in our OM intuitions. He
attempts to formulate a version of consequentialism that will not be over-demanding and integrity-violating. He finds the view that the demands of morality might be incompatible with "a life worth living" to be a "chilling" one, and thus wants to avoid it.17

Toward this end, he offers a distinction between "sector" and "scatter" approaches to morality, based on an analogy with a police squad doing crowd control at a soccer match.18 The best way for a group of people to do crowd control is to divide up, each person taking on a particular "sector" of the crowd for his primary responsibility. If there is an emergency somewhere, individual policemen will (and should) abandon their sectors to help out in another. But overall, the best way to do the job is for everyone to worry about their own sectors. This approach is contrasted with the "scatter" approach, where each keeps trying to work all over the stadium, doing good wherever possible.

Similarly, Jackson argues that each of us should focus our moral attention on the "sector" of those people closest to us, for a variety of reasons. We know those closest to us, and thus know best how to do good to them and secure their welfare.19 We are likely to be able to sustain long, drawn-out courses of action with those closest to us, but probably not with strangers—unless we are of an extraordinary psychological make-up, our persistence in altruistic endeavors for those we barely know is likely to wane fairly quickly, or lead to burn-out.20 We are more likely to be able to see our projects of beneficence through to their conclusions if we focus on

18 Ibid., 473-4.
19 Ibid., 474.
20 Ibid., 481.
those closest to us. Most of us need “close personal relationships” in order to avoid being “outrageously selfish.”

How do we get moderate moral demands out of this? First, those in our “sector,” those we are closest to, are likely to be of roughly the same level of welfare as we are. So, as it turns out, seeing to their welfare probably will not require enormous sacrifice on our parts. Secondly, whatever interest we have in projects of beneficence directed toward those outside of our sector, our ability to carry them out will be morally limited, as we will be forbidden to “abandon” those in our own sector (in the absence of very powerful reasons to think that overall we would do more good ignoring the sector plan.) If my friends and family are living more or less ordinary lives, I must keep up a fairly ordinary life myself if I am to be in a position to keep my commitment to looking out for their welfare. Except in the case of an emergency (which Jackson implies will be quite rare), I cannot drastically sacrifice my well-being for the sake of those remote from me, for that would inhibit my ability to look after my sector.

4. The accounts and supererogation

4.1 Recapitulation

In summary, these four accounts of modest moral demands suggest four views of how moral requirements on individuals are limited:

\[\text{\textit{Ibid.}}\]
Murphy: The requirements of morality are set by the fair share of beneficence to be done.

Williams: The requirements of morality are set by whatever good one can do without jeopardizing one’s independently valued projects and relationships.

Jackson: The requirements of morality are set by whatever maximizes welfare, which will primarily involve taking steps to maximize the welfare of those in one’s “sector” of friends and family and close acquaintances.

Wolf: The requirements of morality are set by whatever good one can do without keeping one from cultivating non-moral tastes, abilities, and interests, in order to allow for an attractive, interesting life.

4.2 Problems

Once we see what views of moral requirements the accounts offer, two problems become evident. I will briefly mention the minor one, and spend most of this section discussing the major one.

The minor problem is this: there is no compelling reason to think that the requirements and demands of morality as described by these theories would correspond with our ordinary intuitive view of morality’s demands. It seems that all of the accounts, with the possible exception of Murphy’s, describe morality’s demands in such a way which makes it quite plausible that they are much higher than we intuitively take them to be.

This is easiest to see with the Williams and Wolf views. It appears that I can give a great deal more than we intuitively think I have to without jeopardizing my ability to have important, independently valuable projects, or a decent, attractive, sociable human life. Consider an
individual trying to decide whether to volunteer for the evening at a soup kitchen or go to the movies, or trying to decide whether to spend one’s money on a camera or donate it to Oxfam. It is difficult to argue that choosing the soup kitchen over the movies, or the camera over Oxfam, would jeopardize the value, meaning, and admirability of an individual life in the way that Williams and Wolf fear. The same goes for giving a large sum to charity in most circumstances. We would have to give a lot, far more than we intuitively think we ought to, in order to threaten our ability to lead a meaningful, valuable, satisfying, sociable life. We might need to choose some sorts of meaningful life over others, but that does not seem particularly problematic, as OM tells us we have to do that anyway. (For example, it seems that according to OM, we ought not to choose the life of a thief or an assassin.)

Jackson’s view also turns out to be far more demanding than OM is. At first, it seems to get us off the hook, by arguing that we should focus on those around us rather than distant sufferers. But the same maximizing consequentialism is present. On Jackson’s view, I still ought to do everything I can to promote the welfare of those in my sector. Now, this requirement is much less demanding than the extremists’, but it is still quite far from the OM standard. We do not ordinarily think we have to do everything we can to maximize utility (or anything else) in the spheres we inhabit. Undoubtedly we ought to be seeking the welfare of our friends and family, those near and dear to us, and this requires doing some things which promote their welfare (and avoiding things which will diminish it). But this stops well short of saying that we ought to do whatever most improves their welfare. Furthermore, Jackson’s view says that should a big enough emergency that you can helpfully address exist outside of your sector, you ought to abandon your sector to address that need. We are permitted to focus on the nearest and dearest only insofar as we can do more good there. Jackson seems successful in arguing that focusing
more on those close to us makes sense, but that “more” looks as if it would be very easily overwhelmed by the immense need out there in the world, while OM still seems to permit us to focus on those close to us even in the presence of such need.

Finally, as was mentioned above, Murphy’s view takes the amount we ought to give to be dependent on the present need and on the ability of others to help out cooperatively in addressing that need. So, its conformity to our intuitions depends on the facts about what the present need is. If Murphy is wrong about the burden of good to be done when equitably distributed, then his view fails to be compatible with OM. However, as previously noted, he seems to have a far better chance of coming close to OM’s view of demands than the others do.

I will not focus on this problem; perhaps those who favor the sorts of accounts described above could flesh them out in such a way as to make them fit more closely with our intuitive judgments. Furthermore, it seems as if those who favor such accounts could argue that at the very least, they come much closer to our intuitions than extremist utilitarian accounts.

A more significant problem with these accounts is that the justification of morality’s modest demands in each case fails to make room for supererogation. This is not to say that the limited-requirement accounts are in every case incompatible with supererogation, but rather that the accounts make it hard to see how supererogation could be accommodated. Perhaps their friends have no interest in doing so, but it seems that if they want to defend our ordinary moral intuitions, they should try to say something about it.

On Jackson’s account, there can be no supererogatory acts. His decision-theoretic consequentialism, while taking into account an agent’s beliefs, subjective probability, avoiding certain risks, etc., is still consequentialist—it says that what the agent ought to do is what has the highest expected moral utility. Even if his sector approach makes the demands of that morality
fairly modest, relatively speaking, as such a “maximizing approach” it is not compatible with supererogation.

On Jackson’s view, normally what will have the highest expected moral utility from my point of view will be attending to my sector—engaging in beneficent acts towards those nearest and dearest to me. When that is the case, those acts are the ones which I ought to do. It turns out to be morally impermissible, in such a situation, to do otherwise. Sometimes, on his view, I will have good reason to believe that my directing my attention outside of my sector will have higher expected moral utility. In those cases, that is what I ought to do—sticking to my sector in such a case will be morally impermissible. Depending on my situation, there will always be some act (or acts) which will maximize expected moral utility from my perspective. This is the only act which is permissible, and in the absence of a tie it is obligatory.

Thus, on this view, the point at which acts stop being obligatory is the point at which they stop producing the maximum good. There can be no supererogatory acts, morally superior acts which are not obligatory. There is no room for “moral saints,” people who consistently do better than they ought. If the sort of approach which Jackson offers doesn’t demand drastic change from ordinary life for a particular individual (after the expected utility has been calculated), voluntarily undertaking such drastic change, even out of compassion and concern, and at great benefit to others, will turn out to be something that individual ought not do.

On Williams’ view, it seems that acts which would put our independently valued projects and relationships in jeopardy would not only fail to be very good, they would be less good than the obligatory acts. For if the value of independently valued projects and relationships is what is doing the work to make it that it is not the case that we ought to do certain acts, it seems that such considerations would dictate that we ought not do them, because they would mean that the
proposed act would violate the agent’s integrity. If violation of the agent’s integrity counts as a consideration against an act’s being obligatory, it also seems to count as a consideration against an act’s being good or permissible.

Williams’ account is not, strictly speaking, incompatible with supererogation. It could be that there are 3 kinds of morally demanding acts: obligatory moral acts (which are generally no more than moderately demanding), supererogatory moral acts (more demanding), and then this group of prohibited or inadvisable acts which would violate the agent’s integrity (too demanding). But the distinction between the obligatory and the supererogatory cannot plausibly be made to coincide with the distinction between the integrity-threatening and the non-integrity-threatening. Another story would need to be told about why some morally good acts are obligatory and some are not.

Wolf allows for people going beyond what she believes morality demands of us, being super-moral, but suggests that this is bad, not good. Presumably we ought to be moral insofar as doing so does not jeopardize our relationships, our ability to be good company and attractive people, our pursuit of a sufficient variety of non-moral interests, etc., but not beyond that point. Again, in this picture, there is no room for supererogation—acts for which it is not the case that we ought to do them, but which are permissible and morally excellent. Morality is obligatory up to the point where it becomes destructively excessive and therefore bad.

Again, like Williams, she could avoid this conclusion. She could introduce a similar three-part distinction: acts which are obligatory, acts which are supererogatory (doing more than we ought, but not enough to threaten our relationships or attractiveness or our pursuit of non-moral interests), and then a third category of impermissible acts which would involve being “too moral”—endangering relationships, attractiveness, and non-moral interests. But then the issue of
what makes some morally good acts obligatory and others not would be left still unexplained, and we are looking for an account which offers such an explanation.

Murphy’s comes closest of all of the accounts examined here to allowing for supererogation. In fact, he states that on his view, any acts which go beyond the level of beneficence that would be required under full participation are supererogatory. Nonetheless, while his account avoids some of the immediate problems which block the others from making sense of supererogation, there are two sorts of difficulties that need to be addressed.

The first is relatively minor, and it regards his views about doing one’s “fair share”. On the one hand, according to OM, it seems that in many circumstances, doing more than one’s fair share is obligatory, especially when the stakes are high. If many children are drowning in Shallow Pond, and you have already saved one, and there are a lot of other people just milling around on the water’s edge who are completely unresponsive to your attempts to get them to help you, it seems that you must go on saving the children. On the other hand, it seems that in other circumstances, doing more than one’s fair share isn’t even good, much less supererogatory. One might argue that it invites moral sloth, and probably constitutes letting oneself be taken advantage of, which is rarely considered a virtue and is usually considered a defect. I think Murphy faces a serious challenge here—he must show that his version of doing more than one’s fair share avoids these two extremes—being good and obligatory, and not being good at all. It seems that there may not be a lot of room between those two alternatives, as the key factor seems to be the seriousness of what is at stake. If it is not very serious, then generally we think you should do no more than your fair share, you should not let yourself be taken advantage of, you should leave room for others to pull their own weight. If it is very serious, then concern with the

---

22 "The Demands of Beneficence," 280, fn. 26
value of standing up for yourself and not letting yourself be used is set aside. But then it will be
difficult to argue that the seriousness of the situation does not make doing more obligatory.

The second problem is that giving more than my fair share does not further the aim of
“promoting the good together with others”; rather, it furthers the aim of “promoting the good”
simpliciter, which is precisely the aim which Murphy wants to argue we do not have. After all,
insofar as we do have that aim, his argument for limited requirements lacks plausibility. But if it
is not a moral aim of individuals to promote the welfare of others themselves, what would make
doing so good? In short, although Murphy explicitly says that going beyond one’s fair share is
on his view supererogatory, what he has to say in justification of these “fair shares” leaves it
murky how this is to be the case. If we consider the aim to promote the good to be an individual
one, then going beyond one’s fair share of beneficence will turn out to be quite good, but it will
be hard to see how it could fail to be obligatory. If we consider the aim to promote the good to
be a cooperative one, then we will have a good explanation for why we are not obligated to do
more than our share of the helping to be done, but it will be far from obvious why it would be
good to do so. Insofar as Murphy is trying to get both limited moral demands and
supererogation, he seems to be sliding between the two views of what makes beneficent acts
good and what our moral aims are.

5. Morality as Hobby?

The accounts of moral requirements we have seen all attempted to argue against the idea
that acting beneficently beyond some point is obligatory. But they tried to do so by offering
reasons which suggest that such acts are not morally good in some respects—at best of neutral or
questionable moral value, at worst morally bad. In trying to argue that such acts are not ones we ought to do, they make it hard to see how it could be good for us to do them, hard to see how there could be any supererogatory acts. They have sacrificed one aspect of ordinary morality in order to protect another.

What emerges is that if we wish to defend OM, we cannot justify modest moral demands by raising doubts about the goodness of doing more than those modest demands require of us. We cannot tie obligation to goodness in that sort of way if we want there to be supererogatory acts. At the very least, it seems that the extremists should be heartened by the difficulty that the defenders of modest moral demands seem to face in coming up with theories which meet their own requirement of fitting with our ordinary moral intuitions.

In closing, I want to suggest that the difficulties with supererogation that the defenders of OM have are not the result of a mere oversight. Offering an account of morality which accommodates our intuitive conception of supererogation is something which looks to be very hard to do.

Part of the problem, I think, is that it is difficult to keep performance of the good-but-not-obligatory acts from seeming like a sort of interest or hobby. Intuitively, we should all learn to develop a certain level of appreciation for the arts, but beyond that, pursuing a deeper knowledge of them and level of performance in them is completely optional. A desirable human life will involve going beyond the minimal requirements in some pursuits and disciplines, but obviously not all—we have to pick and choose. One might think that morality should be treated similarly. We should all pursue being moral to some limited extent, but beyond that it is just one interest/hobby/ability among many which we might choose to cultivate. Thus, when people consider whether or not they should perform the morally better acts available to them, what they
are really asking is whether or not they want to develop this interest in morality further as opposed to other interests. Moral saints are merely masters of their chosen field—comparable to a Yo-Yo Ma, or a Venus Williams. That is the way in which they are unusually good. And their magnificent moral acts are analogous to great performances and achievements in any discipline.

This kind of story turns out to be quite unsatisfactory. One should deny that morality should be treated as an interest or hobby in this sense, for a variety of reasons.

It seems that supererogation is hard to square with some of our attitudes toward moral saints, or those who do the morally better acts. We feel threatened, challenged, convicted by their actions. It seems almost as if we feel we ought to be doing what they do. We generally do not feel the same way about great artists or athletes or scientists. An artist might feel awed or challenged by the performance or the presence of the great artist. But I do not feel that way, because, as far as I’m concerned, the arts are not for me. I am not trying to become like the great artist, and that’s just fine. But in the case of morality, things are different. There is something problematic about saying, “Well, sure she’s amazingly selfless and generous and kind, but that’s just not my thing.” Pursuit of proficiency in music makes one a better musician, but pursuit of proficiency in morality makes one a better person, it seems, on our intuitive view. (Wolf’s remarks to the contrary notwithstanding.)

Furthermore, it is notable that we admire most those moral saints who think they are doing what they ought, or who do not think they have a choice, who could not live with themselves if they neglected to do the supererogatory acts, if they lived only moderately moral lives like our own. The truly admirable moral saints, famously, do not think they are special or doing something exceptional. Those who have performed a striking supererogatory act often say that it is “what anyone would have done” in that situation.
On the surface, this is rather odd, for it seems we admire them more the more false their beliefs about the demands of morality and themselves happen to be. Again, this is very much unlike the artist or the athlete or the scientist. There is a place for humility, but we do not consider it to be a defect in the artist for him to appreciate the greatness of what he has done. In contrast, a moral saint who shares OM’s appraisal of his accomplishment is not much of a moral saint at all. We see the supererogatory acts as good but optional (and we usually opt out of them), but those we respect and admire as morally great see them simply as what must be done.

A third difficulty with the “hobby view” stems from the fact that different disciplines are weighed morally against each other. We value some sorts of proficiency higher than others, in part for moral reasons. Proficiency in pursuits which are morally questionable or bad is not valued very highly at all. And even for pursuits which do not threaten our moral obligations, there is a ranking or evaluation which goes on. Proficiency in the arts has a connection to the meaning and values of human life which, say, proficiency in video-game playing does not, and is thus more valuable. Devoting one’s life to medicine is morally better than devoting it to stamp collecting. But if we evaluate our interests and pursuits in this fashion, how could morality fail to come out on top, as the discipline most worthy of being pursued? What other criteria could we use to rank them besides moral ones, broadly understood? It seems plausible that we ought to thus morally evaluate whatever pursuits and projects we are involved in, and insofar as it is possible without violating our integrity, try to focus our efforts on “higher” ones.

Another question raised by the comparison between morality and hobbies is this: Why should we treat morality on the obligation side of the line so differently from these other pursuits, and treat it as being similar on the supererogation side? Perhaps it is true that a well-rounded human life will have some sort of development in the areas of art, literature, sport,
music, science, philosophy, etc., and so in a sense we might have “obligations” in each of those disciplines. But the obligations of morality are not of the same sort as these. Morality, when its dictates are obligatory, is seen as trumping all other sorts of considerations and obligations, not as one voice among many telling you what you ought to do. Why should it suddenly become “just one of the crowd” beyond a certain limit, stop making demands, and start offering humble suggestions?

In other contexts, “ought” does not have such limited, minimal implications. At a certain point in a chess game, there might be a variety of moves available to me. Two or three of them might be okay. But if one is clearly better than all the rest, then it is the move I ought to make. The same seems true of all other projects and hobbies—within the context of the hobby, what I ought to do is what optimizes my performance. Within the context of the “hobby of morality”, I ought to do the morally best acts.

What is hard to see is how we could ever escape “the context of the hobby of morality”. Isn’t morality intuitively understood to be fundamentally overriding in some sense? It seems that there is something bizarre about weighing the concerns of morality, even beyond the intuitive limits of obligation, against the concerns of aesthetics or sports. If morality does have this sort of overriding, overarching nature, then how could the moral goodness ascribed to supererogatory acts fail to make those acts ones we ought to do? What sorts of reasons could outweigh morality’s recommendations, even if those recommendations are not in themselves obligatory? In short, it seems quite difficult to pry [morally] “good” away from “ought”, and therefore difficult to make sense of supererogation. This is not to say it is impossible, only that it is hard to see how it is to be done.
REFERENCES


Deontological Judgments about Belief

I. Introduction

We often make deontological judgments about acts. We talk about what people ought or ought not to do. We speak of certain acts as permitted, required, or forbidden. We talk about obligations to do certain things. We praise, blame, and criticize people for their acts. In all these ways, we also often make deontological judgments about beliefs.

To see how readily deontological judgments about beliefs come to us, we need only consider W.K. Clifford’s classic story of a ship-owner in his “The Ethics of Belief”:

A SHIPOWNER was about to send to sea an emigrant-ship. He knew that she was old, and not over-well built at the first; that she had seen many seas and climes, and often had needed repairs. Doubts had been suggested to him that possibly she was not seaworthy. These doubts preyed upon his mind and made him unhappy; he thought that perhaps he ought to have her thoroughly overhauled and refitted, even though this should put him to great expense. Before the ship sailed, however, he succeeded in overcoming these melancholy reflections. He said to himself that she had gone safely through so many voyages and weathered so many storms that it was idle to suppose she would not come safely home from this trip also. He would put his trust in Providence, which could hardly fail to protect all these unhappy families that were leaving their fatherland to seek for better times elsewhere. He would dismiss from his mind all ungenerous suspicions about the honesty of builders and contractors. In such ways he acquired a sincere and comfortable conviction that his vessel was thoroughly safe and seaworthy; he watched her departure with a light heart, and benevolent wishes for the success of the exiles in their strange new home that was to be; and he got his insurance-money when she went down in mid-ocean and told no tales. 23

We could make a variety of critical judgments about the ship-owner, and among them are deontological judgments about his belief in the ship’s seaworthiness. He ought not to have believed that the ship was seaworthy. It was impermissible for him to believe it. It was

blameworthy in him to believe it. In this chapter I want to focus on one particular sort of
deontological judgment about belief, those judgments of the form “S ought to believe p” and “S
ought not to believe p.”

Many utterances of the form “S ought to believe p” do not involve deontological
judgments. Suppose, for example, that it were discovered that stimulating a part of the brain
seemed to result in the formation of a particular belief—the belief that pigs can fly.
Experimenters try it out on many subjects, and in every case, when that part of the brain is
stimulated, they end up with the belief that pigs can fly. Then suppose yet another subject,
Smith, goes in, and that part of his brain is stimulated. However, when he is asked about porcine
aviation, it turns out that he finds the whole thing preposterous. The puzzled experimenter might
well say, “That’s odd. You ought to believe that pigs can fly.” In such a case there is no
deontological judgment. None of the other kinds of deontological judgments apply—we cannot
talk about obligations or praise or blame or permissions. Smith does not have an obligation to
believe that pigs can fly. Believing it is not required of him, nor is disbelieving it impermissible.
It would not be right to blame, criticize, or praise Smith for his beliefs. We might discover
something unusual about his brain and “blame” his “failure” to believe on that, but the blame
does not reach up to Smith himself.

How that sort of ought-to-believe sentence is to be understood is an interesting subject,
but outside the scope of this chapter. Here I will be considering sentences of the form “S ought
to believe p” which imply at least some of the other deontological judgments about belief
(DJB’s) mentioned above.

24 I am borrowing the phrase “deontological judgments about belief” from Feldman. I am not quite happy about it,
as certain readers have found calling these judgments “deontological” to misleadingly suggest that there must be
something moral about them. Along with Feldman, I intend to use “deontological” in such a way that makes no
such implication and leaves open what connection to morality these judgments have, if any.
Some philosophers find these deontological judgments about belief worrisome, and among those who are not worried, there is considerable disagreement about what makes it true that “S ought to believe p.” In the second section of this chapter, I examine the main problem that has been raised regarding DJB’s. Briefly, the problem is that three things seem plausible, and they are jointly incompatible:

1) Deontological judgments about beliefs are sometimes true.
2) We do not have control over our beliefs.
3) If a deontological judgment is true, we have control over whatever is being judged.

I assume (1) in this chapter, and so do all the philosophers whose views I discuss here. I will examine two attempts to show that (2) is false (Matthias Steup and Sharon Ryan) in the third section of this chapter, and two attempts to show that (3) is false (Nicholas Wolterstorff and Richard Feldman) in the fourth section. Ultimately, I find their proffered solutions wanting.

I then go on in the fifth section to discuss one particular proposal about what makes DJB’s true—Feldman’s as expressed in his “The Ethics of Belief”—try to make some trouble for it, and offer an alternative of my own. I argue that we do have a kind of control which supports our making deontological judgments about beliefs, but such control is very different from the sort depicted by Steup and Ryan. In the sixth part, I consider and address some possible objections to my view.

It is worth noting that I am approaching this question from a very different direction from that of most of the people (i.e., epistemologists) who address it. I am interested in the concepts ‘ought’, ‘obligation’, and ‘deontological judgment’, and am interested in our everyday deontological judgments about belief because of that. Most writing on this subject has been done by those who have a favorite theory of knowledge or epistemic justification, and want to marshal deontological talk about beliefs to support it. I do not want to say at the outset that such a view
is wrong; I do, however, want to suggest that we should be open to the possibility that our ordinary talk of what we ought to believe might not neatly map onto what the best theory of knowledge says that justification is.

Clifford’s ship-owner story shows that moral and epistemic issues can be closely intertwined. The ship-owner’s fault is not merely epistemic. If he had instead cherished an unjustified belief about a matter of much less consequence—say, whether or not a particular chair was sturdy enough to hold him, we would still say he ought not to believe that the chair will hold him. But there would be something far less heinous about it, as the moral significance of his belief would be greatly reduced. Cases of racist or sexist belief, among others, also elicit both epistemic and moral judgments. For simplicity’s sake, in this chapter I will focus primarily on examples which do not raise moral issues, but virtually all of what I say applies to morally impermissible belief as well.

2. The Problem with DJB’s

The classic reason for thinking that talk about what we ought to believe is problematic is as follows: Deontological judgments about beliefs cannot be true, as they require voluntary control, which we do not have over our beliefs. We cannot, it is pointed out, believe at will or believe voluntarily. It is not up to us in the right sort of way.

Bernard Williams’ “Deciding to Believe” thus argues that precisely because belief aims at the truth, our beliefs cannot be under our control. Believing is not like blushing, which is involuntary but only contingently so.25

If I could acquire a belief at will, I could acquire it whether it was true or not. If in full consciousness I could will to acquire a ‘belief’ irrespective of its truth, it is unclear that before the event I could seriously think of it as a belief, i.e. as something purporting to represent reality.\textsuperscript{26}

Not only do we know from experience that we cannot believe at will—there is something dubious about the very idea of it, because believing involves taking something to be true. Suppose I could come up with beliefs at will---the belief, for example, that it is raining outside. But if I pause for a moment to reflect, I will remember that my belief was caused by my simply willing it to be so, and that is clearly no reason to think it to be true. Believing at will would sever the connection between belief and truth, and thereby would make belief something else altogether.

Once this much is granted, the traditional doctrine of “ought implies can” can be relied upon to create a problem for DJB’s. Since it is unclear what “can” amounts to here, a better way to put the doctrine in this case is “Ought implies control,” where this could be put:

\[(O\rightarrow C): \text{If S ought to } \varphi, \varphi-\text{ing is something that S has control over.}\]

Of course, what “control” amounts to can also be quite complicated, depending on one’s view of free will. But something like \((O\rightarrow C)\) does seem to be in the background of our judgments about what S ought to do. It explains why hard determinism is thought to threaten morality, and why compatibilists feel the need to tell some story about how we can have control in some significant sense over what we do even if determinism is true.

We can put the puzzle this way: On the one hand, it appears that deontological judgments imply control, and it appears that we do not have voluntary control over our beliefs. On the other hand, we have a common practice of making deontological judgments about beliefs.\textsuperscript{26} Ibid.
How should we deal with this apparent conflict? For those who do not want to reject the possibility that DJB’s can sometimes be true, there are two alternatives. The first (which I discuss in section 3) is to argue that we do have the appropriate sort of control over our beliefs. The second (which I discuss in section 4) is to argue that even if we do not have such control, control is not necessary for DJB’s, that is, (O→C) is not true for beliefs. I will discuss two arguments of each type.

3. Arguments that we have control over our beliefs

3.1. Steup’s Doxastic Decisions

Matthias Steup argues that we can have the right sort of control over our beliefs, because beliefs can emerge from what he calls doxastic decisions. Just as we have control over our actions if we carry them out because we have decided to do so, we have control over our beliefs if we hold them because we have decided to do so.

This idea of doxastic decisions is somewhat unintuitive, so Steup tries to make it more plausible by appealing to a comparison between practical deliberation and epistemic deliberation. In practical deliberation about what to do, we consider reasons for and against different actions, and these considerations can result in a decision to do something. Steup suggests that similarly, in deliberation about what to believe, we consider reasons for and against different doxastic attitudes toward propositions, and these considerations can result in a decision to believe something—a doxastic decision.

The criticisms of Steup’s view I have seen assume that he considers deliberation to be necessary for doxastic decisions. Sharon Ryan, for example, objects that we appear to come by

---

28 Ibid., 33
many of our beliefs without any deliberation whatsoever. If deliberation is necessary for doxastic decisions and voluntariness, then if Ryan is correct, Steup has secured voluntariness for only a small subset of our beliefs. He clearly wants to accomplish more than this with his argument. Richard Feldman offers a regress argument (mentioned with approval by Ryan)—if deliberation is a matter of considering our other beliefs, not all beliefs can be a product of deliberation. At some lower level, the beliefs must not be formed by deliberation, and therefore involuntary, and Feldman seems to suggest that this would threaten the voluntariness of the "higher" beliefs founded on them. (The argument is not made out in any detail.)

But these criticisms do not touch Steup’s view, as he points out in a footnote: “I am offering only a sufficient condition for the execution of a doxastic decision, for it seems to me that doxastic decisions can be executed as well in ways that do not involve prior deliberation about one’s evidence.” Given his reliance on the analogy between decisions to act and what he calls “doxastic decisions,” this position makes sense. We often decide to do things without deliberation. If belief is to be treated similarly, why think that we need to deliberate in order to decide to believe?

Steup’s key claim is that the processes of forming a belief and deciding to act are sufficiently similar so that saying we have control over the latter but not the former would be arbitrary. He tries to show that we have no good reason to think that they are different. Toward this end, he argues plausibly that a variety of alleged differences do not hold between all cases of putative decisions to believe and all cases of decisions to act. For example, to the objector who

---

30 “Doxastic Compatibilism,” 69. The criticism comes from written comments Feldman made on an earlier draft of her paper.
31 “Doxastic Voluntarism,” 34.
claims that decisions to act require effort and time to execute them, while decisions to believe do not, Steup points out that some decisions to act do not require much effort or time at all, such as deciding to think about one’s last summer vacation. Furthermore, Steup claims, some decisions to believe do require time and effort to execute. A racist might come to realize that his beliefs are unsupported by the evidence and must be abandoned, but yet have difficulty in doing so, since “old beliefs may be hard to shake off.”

In spite of these efforts, Steup misses an important difference. In the case of epistemic deliberation, even at the very start, I hold some doxastic attitude, some degree of belief or disbelief, toward the proposition in question. As I weigh each of the considerations, I may shift towards greater or lesser degrees of belief or disbelief. But my “decision” is simply the state I am in when I stop weighing considerations. There is no real decision to believe at all. I may decide that I have reflected enough on the matter, but this only grants an air of finality or stability to a belief that already exists.

This is different from an objection Steup considers, that after a doxastic decision is made, there is nothing left to do (and therefore doxastic decisions are significantly unlike decisions to act.) The objection here is that before the decision is made, there is nothing left to do. As I consider the evidence, my doxastic attitude toward p, my belief or disbelief in p, is automatically responsive to each bit of data. Steup himself acknowledges “Sometimes one tends to have a view on the matter already while sifting through the evidence.” What I question is the idea that one sometimes does not.

---

32 “Doxastic Voluntarism,” 35.
33 Ibid.
34 Ibid., 34, note 14.
This is quite unlike deciding to act. In deliberations about what to do, one decides and acts. In these “epistemic deliberations,” the believing, the holding of some doxastic attitude or other toward the proposition in question, occurs throughout the deliberation. Furthermore, in deliberation about what to do, the decision draws deliberation to a close. I weigh my considerations, and then I decide, so I can go do whatever it is that I decide. But do we need to draw epistemic deliberations to a close with a doxastic decision? I would say no. Perhaps I decide to stop deliberating, because I run out of evidence to consider, or perhaps because I want to do something else, but that is not a decision to believe.

A common view of practical deliberation provides us with another reason to reject Steup’s proposal. On this view, epistemic deliberation is part of practical deliberation. In ordinary practical deliberation, we first deliberate epistemically, until we settle on a belief about what would be best to do. This is followed by a practical component of the deliberation, which results in an intention to act. So practical deliberation and epistemic deliberation cannot be treated as similar, because the former includes the latter plus something else. If epistemic deliberation culminated in a decision, then practical deliberation would involve two decisions—a decision about what ought to be done, and a decision to do it. I do not wish to argue for this view of practical deliberation here, although I find it quite plausible. I mention it only to show that those who are drawn to that view will have another reason to think that Steup’s epistemic deliberations are not as much like practical deliberations as he suggests.

Given these differences between deliberating about what to do and deliberating about what to believe, I suggest that we should be reluctant to accept Steup’s suggestion that we should treat them in the same manner.
He muddies matters by taking for his paradigmatic example of epistemic deliberation and doxastic decision the case of a man trying to assess whether O.J. Simpson was guilty or innocent.\textsuperscript{35} This sort of example appears to support his theory by conjuring up images of jury deliberations and verdicts. But a jury deliberation is aimed at a decision about what to do (acquit or convict), not a decision about what to believe.

Perhaps he could appeal to the fact that we say things like “I don’t know what I ought to believe about OJ’s guilt.” Initially, such expressions do make it sound as though a decision is in order, sounding similar to “I don’t know what to wear.” But the superficial similarity is misleading. In most cases where we say “I don’t know what I ought to believe about ______,” we do know what doxastic attitude we ought to hold at the present moment, given our current evidence—we ought to remain neutral, as our evidence is either too limited or too equivocal. We are saying that we do not know the truth of the matter. The solution to that plight is not a decision, but further information-gathering.

3.2. Ryan’s intentional believing

While Steup tries to convince us that we have control over our beliefs because there is no significant difference between doxastic and practical decisions, Sharon Ryan claims that we have the right sort of control because we believe intentionally. She challenges Williams directly: while he compared believing to blushing, she claims that “the comparison has no merit.”\textsuperscript{36}

One might, for some strange reason, walk into a CD store proclaiming, “I’m not going to believe there are any CDs in the store”. If this person opens her eyes, she will fail. However, if she’s normal, once she opens

\textsuperscript{35} “Doxastic Voluntarism,” 34.

\textsuperscript{36} “Doxastic Compatibilism,” 72.
her eyes, she will find herself with a new intention. Once she sees the CDs, she will intentionally believe there are CDs in the store.\textsuperscript{37}

Elsewhere she says, "When you believe a proposition, you’re doing something…”\textsuperscript{38} and “I’m believing freely because I’m believing what I mean to believe.”\textsuperscript{39} She offers little, as far as I can tell, in the way of argument for this intuitively baffling view. So why should we agree?

At bottom, her reasoning follows lines similar to Steup’s. On their view, an adequate case has not been made for treating beliefs differently from ordinary voluntary actions. Ryan’s target is the claim that belief and action differ in that we can “just” decide to act while we cannot “just” decide to believe. Those who argue that belief is involuntary tend to compare beliefs with actions produced by a decision about which the agent was more or less indifferent. I can just decide to put up my hand, or put it down—I do not need much in the way of reasons for doing so; I could just as well not do so. I can just decide to order one entrée rather than another at a restaurant, without any compelling reason to favor the one over the others; I could just as well pick another. It is true that cases of belief are not like this. But Ryan rightly points out that there are plenty of uncontroversially voluntary things that we cannot “just” decide to do in that sort of way. Sometimes we have powerfully compelling, decisive reasons to choose one action over the alternatives, but we do not think that this diminishes the voluntary nature of the action. For example, if offered a choice between fresh pumpkin soup and “one month old lima bean and raw calf’s liver stew,” Ryan finds that her reasons compel her to pick the pumpkin soup.\textsuperscript{40}

Nonetheless, it seems her choice of the pumpkin soup is voluntary. If that were the only reason for worrying about voluntary control for beliefs, we should be persuaded.

\textsuperscript{37} Ibid.
\textsuperscript{38} Ibid., 73.
\textsuperscript{39} Ibid., 74.
\textsuperscript{40} Ibid., 73.
But it is not the only reason, and we might well still worry. Ryan focuses on relatively unusual examples of doing something or choosing to do something, and perhaps believing is like these cases of doing something, at least in some respects. But it is reasonable to ask whether we would make deontological judgments about acts if all acts were relevantly similar to the soup case. Examples like that are on the “fringe” of the realm of voluntary action. Suggesting that cases of belief are similar to these is thus not enough to show that believing is voluntary or intentional. Many voluntary actions are nothing like the soup case—there is such a thing as acting on a whim. When the stakes are high it is true that I cannot “just” choose to do something, but usually the stakes are not so high.

In contrast, there is no such thing as believing on a whim. Regardless of how low the stakes are, how insignificant the belief is, we cannot just decide to believe something. This difference seems significant. Steup attempts to explain it away by saying that “with beliefs, there always is something important as stake: truth.” Every case of belief is like the soup case. But this explanation seems inadequate: why are all truths so important? On the face of it, it seems that there are many truths that do not matter much to me at all. The exact number of people in Luxembourg, for example, is of no concern to me. So why can I not form a belief about it on a whim?

Furthermore, beliefs are significantly different even from Ryan’s “fringe” cases of action. It is not at all clear that she had to have the pumpkin soup—she could have refused both the soup and the stew had she felt like it. Even if she were starving and could not resist the soup, she could choose how to eat it, whether to guzzle it or sip it. There are all sorts of control which we have over our doings which we do not have over our believings.

---

41 “Doxastic Voluntarism,” 33.
4. Arguments that try to show that control is not required for DJB

If the prospects are dim for establishing that we have control over our beliefs, an alternative is to argue that control over beliefs is not necessary for deontological judgments about them to be true—that is, that \( O \rightarrow C \) is false, at least when it comes to beliefs. The fact that we ought to believe something does not imply that we have control over our beliefs.

Such an approach may be inspired by the work of Michael Stocker\(^{42}\) and Walter Sinnott-Armstrong\(^{43}\), who argue that when talking about morality, ought does not always imply can. (Sinnott-Armstrong’s conclusion is that ought only conversationally implies can, in the Gricean sort of way. If we are concerned about defending our ordinary deontological judgments about belief, that might be bad enough for us.)

The most plausible cases each has to offer are those of culpable inability—making a promise you know you cannot keep, or making a promise you can keep and then rendering yourself unable to keep it. But we need not restrict the point to promises. Consider a star player on a soccer team, who before an important match I does something foolish and reckless, injuring herself so that she cannot play. Someone could say to her reproachfully, “You ought to be playing today.” Afterwards, when her team loses because of her absence, they can snarl at her, “You ought to have played today.”

Neither Stocker nor Sinnott-Armstrong takes himself to show that our ought-judgments have nothing to do with ability; they simply argue that it is not true in every case that you must have control over the thing that you ought or ought not to do. Similarly, defenders of


deontological judgments about belief who take this route do not merely take it for granted that (O→C) is false in these cases. They think they need to give some sort of explanation of why in these cases, "ought" judgments do not require control. In this section of the chapter I will discuss two such explanations, one offered by Nicholas Wolterstorff, the other by Richard Feldman.

4.1. Wolterstorff's Paradigm Obligations

In the spirit of reconciling those who accept the control thesis (O→C) and those who reject it, Nicholas Wolterstorff offers a solution: that "ought to believe" statements express what he calls "paradigm obligations". On his view, there are two kinds of obligations: responsibility obligations, and paradigm obligations. Responsibility obligations are "connected with accountability," with praiseworthiness and blameworthiness. Paradigm obligations, on the other hand, have nothing to do with accountability, praise, or blame, and are expressed by statements like "You ought to be walking on it in about a week", "It should have black spots all over its shell, not just around the rim", and "That's strange; you ought to be seeing double."

Wolterstorff has little to say by way of defining this sort of obligation, he mainly gestures toward it with examples. He does say that it has something to do with "proper functioning" as opposed to "malformed functioning."

(O→C), Wolterstorff claims, is true for the responsibility obligations, but not for the paradigm obligations, and our epistemic obligations are paradigm obligations. Thus, on his view, even though (O→C) is false for beliefs, we can still have obligations to believe, because they are of the paradigm variety.

---

Feldman’s reply to Wolterstorff, which seems correct to me, is that these “paradigm obligations” are not obligations at all. They might involve the word “ought”, but they do not express obligations or any sort of deontological judgment. Wolterstorff says, as noted above, that the paradigm obligations have nothing to do with accountability, praise or blame, but those things are very much a part of our ordinary DJB’s. Sometimes we do hold people responsible for their beliefs, as in the case of Clifford’s shipowner. This proposal cannot help us defend those judgments.

Furthermore, it is unclear what Wolterstorff’s paradigm obligations are supposed to be. He offers various “ought” judgments where responsibility is not implied, and treats them as being the same, but on further examination they are quite different. The first case (“You ought to be walking on it in about a week”) has to do with proper function—if the healing processes in the leg, and the treatment you have been given, work right, you will be walking on it in about a week. The second case (“It ought to have black spots all over its shell”) has to do with something being a typical specimen of its kind—if it were a typical ladybug, it would have black spots all over its shell. The third case (“You ought to be seeing double”) involves the “ought of expectancy:” if things had gone as expected, you would have seen double. These different sorts of “ought” judgments can come apart. Consider the “typical case” of a particular illness, which by definition involves a defect in the functioning of the body: “X ought to have itchy red patches all over his body tomorrow.” In one sense X ought to have itchy red patches all over his body, because that is how this disease typically progresses, that is to be expected. Yet in another sense,

---

46 Richard Feldman, “The Ethics of Belief,” in Evidentialism, Earl Conee and Richard Feldman, (Oxford: Oxford University Press, 2004), 174. Note that Wolterstorff’s paradigm obligations are just like the case of Smith who fails to believe that pigs can fly, discussed at the beginning of this chapter.
X ought not to have itchy red patches all over his body. You can see that he is sick because he ought not to have such patches (if he were healthy, he wouldn’t) and yet he has them.

Thus, Wolterstorff's paradigm-obligation account faces two difficulties. Not only does it fail to get us epistemic obligations and other DJB’s, it is not even clear what these paradigm obligations are, what unites them besides the fact that no responsibility is involved.

4.2. Feldman’s Role Obligations

Feldman’s solution to the puzzle of how DJB’s can be true is that obligations to believe are what he calls role obligations. Role obligations come with a particular role, and do not depend on one’s ability to perform them. A teacher ought to explain things clearly, in virtue of being a teacher, even if he is utterly incompetent and cannot do so. “Parents ought to care for their kids. Cyclists ought to move in various ways.” Feldman argues that believing is these cases.

We should be wary of embracing this view, since as Ryan points out, believing is hardly a role we take up. I want to consider a little more carefully, though, some of the relevant differences between believing and Feldman’s examples.

First, it is plausible that the teacher’s obligation to explain things clearly stems from the voluntary nature of his taking up the role in the first place. We might well wonder why he took up this line of work anyway if he is so incompetent. In the case of the parents, we might wonder why they decided to have and keep a child if they could not care for it adequately. Such people have done something similar to making a promise they cannot keep. It is no longer intuitively obvious that the teacher has that obligation, say, if he were impressed into teaching at gunpoint,

---

47 "The Ethics of Belief," 175.
48 "Doxastic Compatibilism," 61.
or required to teach something he knew to be false. But being a believer is not a role that we voluntarily take up—we cannot escape it, we cannot choose otherwise. If voluntary acceptance of a role is necessary for role obligations to hold, role obligations to believe are suspect.

Second, the cases are importantly dissimilar, because the incompetence in the teacher case is nothing like the lack of control-aptness in the belief case. The teacher can make efforts, however pathetic, in the direction of teaching well. He can try and fail. Belief is not like that, for we do not even know what it is to try. The two inabilities are altogether different.

Finally, as with Wolterstorff's examples, it is unclear that Feldman's three examples all illustrate the same kind of thing—the same sort of "role obligation". To my ear, "Cyclists ought to move in various ways" sounds more like "Clocks ought to tell correct time" than "Parents ought to care for their children." (Not quite the same, but closer.) Are cyclists to blame for their failure to move in the ways that count as good cycling? Does it make sense to hold them responsible? Do they have an obligation to move in such ways? One awkward example does not invalidate the view, of course, but since the account of what the "role obligations" are is little more than a gesturing toward a handful of examples, the fact that one of the examples differs wildly from the others makes it hard to know what the "role obligations" are supposed to be.

For all these reasons, Feldman's "role obligation" view leaves much to be desired as an account how we can make DJB's given that our control over our beliefs seems doubtful.

5. What makes it true that S ought to believe that p?

Up to this point we have been looking at a problem for those who want to defend our deontological judgments about beliefs. As discussed above, the problem is that three things seem plausible, and they cannot all be true:
1) Deontological judgments about beliefs are sometimes true.
2) We do not have control over our beliefs.
3) If a deontological judgment is true, we have control over whatever is being judged.

First we looked at two accounts which attempt to show that we do have control over our beliefs in the right sort of way—Steup’s “doxastic decisions” and Ryan’s “intentional believing.” Then we looked at two accounts which try to show that our DJB’s can be true even though we do not have control over our beliefs—Wolterstorff’s “paradigm obligation” account and Feldman’s “role obligation” account. None of these efforts look promising.

I will set aside the problem for the time being, and suppose that the deontological judgments about belief we make are sometimes true, and consider what makes them true. Considering what makes them true, I hope to show, helpfully points to a way of dealing with the problem.

The best worked-out account of what makes DJB’s true appears to be Feldman’s. In his “The Ethics of Belief,” after arriving at his role-obligation solution to the control problem, he offers and defends a view about what makes DJB’s true:

> For any person S, time t, and proposition p, if S has any doxastic attitude at all toward p at t and S’s evidence at t supports p, then S epistemically ought to have the attitude toward p supported by S’s evidence at t.49

On his view, what one ought to believe at t is dependent on the totality of the evidence that one has at t, but what counts as evidence is somewhat complicated. He does not address the question in “The Ethics of Belief,” but has something to say about it in his earlier “Having Evidence,” where he says that beliefs must meet an availability condition and an epistemic

49 “The Ethics of Belief,” 178.
acceptability condition in order to count as evidence. The availability condition he arrives at is:

"S has p available as evidence iff S is currently thinking of p."

He does not discuss the epistemic acceptability condition in any detail, except to say that it should rule out beliefs which we have "for no good reason", and that "Only things that I believe (or could believe) rationally, or perhaps, with justification, count as part of the evidence I have."

This is a classic example of the sort of approach I mentioned earlier in the chapter, as Feldman starts with a robust view about justification, an evidentialist view, and works out a view about DJB's from there. I will argue in this section that his view does not capture many of our ordinary ought-to-believe judgments.

Consider an abbreviated, modified version of an example offered by Feldman in his "Epistemic Obligations":

Suppose an unfortunate student, Jones, is about to take an oral exam with an unusually difficult teacher. In fact, the failure rate on oral exams with this teacher is grim indeed: only 10% of the students who take the exam pass it...Let us also suppose that Jones, the student, has no particularly good information about his abilities on the material to be covered on the tests.

It seems clear that Jones ought not to believe that he will pass. But this story just gives us the bare bones, and there are various ways the details could be filled out. Consider three ways in which Jones might come to believe that he will pass.

- FLATLY IRRATIONAL: Jones believes that the pass rate is 10% and that he has no good information about his own abilities, and nonetheless believes that he will pass.

- WISHFUL THINKING: Jones learns that the pass rate is 10%, and realizes that he has no good information about his own abilities. But, disturbed by the likelihood of failure, he begins...
He begins to doubt whatever source told him that the pass rate was 10%—it couldn’t really be that low, so-and-so might be untrustworthy, perhaps there was a misprint. He dwells on the fact that the teacher seems to genuinely like him, while ignoring the fact that the teacher also seems to be extremely fair. He reformulates his view of his abilities—perhaps taking his performance on an altogether different sort of test to be relevant after all. And thus he comes to believe that he will pass.

EVIDENCE AVOIDING: Jones is vaguely aware that the teacher is difficult, but his fear of contemplating the possibility of failure leads him to avoid realizing how difficult the teacher really is. The pass rates for the exam are listed on a sheet passed out in class, but he specifically avoids looking at them and thus does not know what they are. He also avoids thinking about his particular lack of abilities regarding the material covered in the exam. Instead, he focuses on past successes in other classes (which, independently of the specific information available to him about this class, would be good evidence) and believes that he will pass.

Feldman’s account picks out FLATLY IRRATIONAL as a case in which Jones ought not to believe that he will pass. What I want to argue here is that WISHFUL THINKING and EVIDENCE AVOIDING (which are far more common than cases like FLATLY IRRATIONAL) are cases where we might well say that Jones ought not to believe that he will pass, and Feldman will have a hard time explaining why.

This is because Feldman’s account makes believing impermissibly strictly a matter of believing in a way that flatly “flies in the face” of the evidence that one has. But what we often have in mind when we criticize people for not believing what they ought is a more complicated matter.

In WISHFUL THINKING, Jones senses a conflict between his evidential beliefs and what he wants to believe, and responds by adjusting his evidential beliefs so that they no longer conflict with what he wants to believe. He cannot do this willy-nilly by deciding to believe different things. It happens more subtly—he over-scrutinizes sources which tell him what he does not want to hear, and thus finds reasons to doubt them. He under-scrutinizes sources which tell him what he wants to hear, and thus fails to find reasons to doubt them. If this goes on long

---

54 Recall the ship-owner case at the beginning of this chapter. (30)
enough, and especially if with the passing of time he forgets (or fails to attend to) exactly how he came to believe what he has in this way, he may find himself with a set of beliefs that support his desired conclusion. Jones’ may be an extreme case of this sort of epistemically defective thinking, but something like it is quite common to a lesser degree.

In EVIDENCE AVOIDING, Jones avoids evidence that would conflict with what he wants to believe, and relies only on evidence that supports it. Thus, his evidential beliefs do support what he wants to believe. He might have an uneasy feeling about what he might learn if he considered other evidence that he could gain (which motivates his evidence-avoidance in the first place), but that by itself does not count for much in the scales of the evidence.

Thus, if I say “S ought not to believe p,” I do not think that all that can be meant is that S’s evidential beliefs do not support that belief. It could be that S has gotten himself into such an epistemic predicament that given everything else he believes, his evidential beliefs do support p. In fact, that seems to be the case with most instances of epistemically culpable belief. There may very well be no immediate defect (at t) of rationality present—and this is what Feldman thinks is required for it to be the case that S ought not to believe p. People might believe what they ought not to believe or fail to believe what they ought because they have ignored heaps of evidence, failed to consider other perspectives, engaged in a variety of forms of self-deception, and so forth.

Even if we were prone to irrationalities of the sort that Feldman describes, what began as a Feldman-forbidden-belief would likely evolve into something more. One who started out flatly-irrationally believing that a certain political or religious leader or group is always right would comes to watch, read, and listen to sources which support his perspective and associate with like-minded people who reinforce his views. He might also come to believe that those
sources which present a different view are offering propaganda and lies, and that those who believe them are hopelessly deceived and not worth listening to. Thus, the evidence input would be changed, and he would eventually work himself into a state where his beliefs are supported by the evidential beliefs (and the evidence) with which he has equipped himself. We can paint ourselves into an epistemic corner, so to speak.

Believing directly against your evidential beliefs, i.e., believing Feldman-impermissibly, is far too difficult for anyone without severe and unusual cognitive defects to account for most of the beliefs we make negative deontological judgments about. It is far easier and more common to influence our evidential beliefs in various ways. Most of us cannot believe Feldman-impermissibly, which is precisely why we get involved in things like EVIDENCE AVOIDING and WISHFUL THINKING. These indirect methods of manipulating our evidential beliefs or our available evidence exist because believing directly against our evidential beliefs is impossible for most of us.

We simply know too little about the evidential states and inner epistemic lives of others to make DJB’s, if those judgments are about what is supported by the evidential beliefs that the person has. It is rare to have such information when making our everyday judgments, unlike philosophical examples where it is built into the description of the situation. But we often know what information is readily available to all, so plainly in view that it would take great deal of effort to not to have it as part of one’s evidential beliefs. When we say that adults in our society ought to believe that the earth is round, or that they ought not believe that we are being visited by extraterrestrials, or that they ought to believe that Elvis is dead, we do not need to know much about the evidential beliefs that they have. We assume that if they live in society with us, there is no way they could believe those things without some sort of self-deceptive practice. We do
not add a proviso to our DJB’s such as, “S ought to not believe that the earth is flat—unless S’s current evidential beliefs indicate otherwise.”

To return to the discussion in earlier sections of this chapter, it is worth noting that for these sorts of cases, which are the most common ones, there is a way out of the control puzzle. Specifically, we are holding people responsible in each case for things they do have some degree of control over—self-deceptive practices such as avoiding evidence or wishful thinking—not the beliefs themselves, but the physical or mental acts responsible for influencing those beliefs. We do not need to find a sort of deontological judgment that does not require control; nor do we need to try to show that beliefs are held intentionally or voluntarily. \((O \rightarrow C)\) does not exactly hold, but we can see why something like it is true.

Furthermore, it is these cases rather than Feldman’s which call for deontological judgments. The person who believes Feldman-impermissibly, like Jones who believes all the evidence against \(p\) yet somehow believes \(p\) nonetheless, is not an appropriate target for DJB’s. Such cases are more like Wolterstorff’s paradigm obligations, or the ought of expectancy. Jones is not to blame for his belief—he cannot be held responsible for it. There is simply something wrong with his head. Consider, by way of contrast, the epistemic culpability of George Washington, who reportedly said of slavery “I shall frankly declare to you that I do not like even to think, much less talk, of it.”\(^{55}\) He avoided (or at least tried to avoid) concluding that slavery is wrong by trying not to think about various aspects of slavery and the humanity of his slaves. He is a proper target of epistemic blame in a way that a man who accepted all the relevant evidence but out of sheer stupidity could not see that slavery is wrong would not be.

6. Objections

I have argued that in most ordinary cases of DJB's, the person's evidential beliefs actually support his conclusion. Let us consider a few objections to this view.

6.1. Objection 1: The alleged evidential beliefs do not meet the standards for evidential beliefs.

According to this objection, the supposed counterexamples (EVIDENCE AVOIDING and WISHFUL THINKING) are not cases of a person believing what his evidential beliefs supports, because the alleged evidential beliefs fail to meet the epistemic acceptability condition. If someone engaged in such self-deceptive practices, the resulting would not be sufficiently rational or justified to count as evidential beliefs. Feldman, after all, is not advancing a mere coherence theory. One's beliefs have to have some positive epistemic qualities to count as evidential.

My response is twofold. First, it is sufficient to make trouble for Feldman to cast doubt on whether our negative DJB's are always made against those who are clearly believing against their evidence. Even if WISHFUL THINKING and EVIDENCE AVOIDING were not clear cases of a person believing what his evidential beliefs support, they nonetheless appear to be cases where a person is not believing something contradictory to his evidential beliefs yet is still believing as he ought not.

But I would go further than that. I think the products of wishful thinking and evidence-avoiding would pass Feldman's acceptability test, although it is hard to tell given how little he says about it. This is clearest in the case of evidence-avoiding. There is nothing wrong with the evidential beliefs the person has. Their quality, justification, and rational support are
impeccable. The problem is that he has avoided obtaining a lot of other quality evidence. It does not seem that this should bear on the quality of the evidential beliefs he possesses.

Now consider wishful thinking. I suggested that we might do this by over-scrutinizing some beliefs and sources and under-scrutinizing others. Undoubtedly there is a degree of under-scrutiny (and perhaps over-scrutiny as well) that would absolutely render acceptance or rejection of a piece of evidence as irrational and unjustified. But there is a wide range of permissible degrees of scrutiny and belief-examination. We allow for personality differences in the matter—some people have a more skeptical nature, some are more epistemically easy-going. The amount of scrutiny we apply to any particular belief is often a matter of what our interests and concerns are at the time. Sometimes we want to be certain, sometimes it does not matter as much.

I would suggest that there are instances of wishful thinking that use only degrees of overscrutiny and underscrutiny within the acceptable range. It might very well be that each particular evidential belief considered on its own merits meets the epistemic acceptability criterion. Where we go wrong is not in our particular doxastic attitude toward any bit of evidence, but rather in our “unfairly” inconsistent handling of different pieces of evidence—putting some bits through the wringer, while letting others off easy.

6.2. Objection 2: These aren’t epistemic oughts.

Feldman considers an objection to his view from hypothetical challengers who think that things like negligence in avoiding evidence or a duty to gather evidence figure into deontological judgments about belief. Although the suggestion here is not the same as the ones he considers,\(^{56}\)

\(^{56}\) Feldman’s hypothetical objectors argue that there is a duty to gather evidence behind our DJB’s, and that those who avoid evidence or do not pursue it with enough vigor are therefore epistemically culpable. While I consider
it is similar enough that he might treat it in the same way. His response is that he is only
interested in what he calls the “epistemic ought.” All these other “oughts” that might have to do
with avoiding evidence or gathering it are prudential or moral in nature. 57

The reply is odd, for even if we accept a variety of different oughts, it is hard to see why
we should think these particular cases are prudential or moral. Feldman does not offer us any
reason why other than that they do not fit in his view of the “epistemic ought”, and thus need to
be accommodated elsewhere. (For him, “epistemic ought”, “moral ought”, and “prudential
ought” appear to exhaust the options.)

But it appears there is an epistemic reading (or at least, a clearly non-prudential and non-
moral reading) of “Jones ought not to believe that he will pass,” even in the cases above where
Jones is engaged in some form of self-deceptive practice. There is no reason to think that
matters of prudence or morality must be at stake. It is true that I might think that Jones on moral
grounds ought not to believe that he will pass, perhaps that he owes it to his fellow man to give
up his irrationality. (Clifford seemed to have something like this in mind.) I might think that for
reasons of prudence he ought not to believe it, that he would be better off if he faced the facts.
But I need not have either of these in mind in order to believe that “Jones ought not to believe
that he will pass.” His believing that he will pass is deserving of criticism on epistemic grounds,
because his self-deceptive practices effectively break the connection between world, evidence,
and belief that allows one’s beliefs to accurately reflect the facts, that allows one’s beliefs to
“aim at the truth.”

---

57 The Ethics of Belief,” 189.
6.3. Objection 3: The purported counterexamples are not counterexamples, because Jones’ evidence in those cases supports the conclusion that he will fail.

One might say that the counterexamples (WISHFUL THINKING and EVIDENCE AVOIDING) are not really counterexamples to Feldman’s view. In both cases, it seems Jones must know the conclusion he is trying to avoid in order to avoid it successfully. He must have good reason to think it true in order to know to start the work of self-deception. He must know something about the reasons for thinking it true to avoid encountering them or to undermine them via a process of wishful thinking. Taking this into account, the objection goes, the evidential beliefs he ends up with will not really support the belief that he will pass.

After all, in EVIDENCE AVOIDING, why does Jones avoid looking at the handout if he does not already believe that it had bad news on it? In WISHFUL THINKING, won’t he remember that he once believed that he would fail, or that his old evidence supported that conclusion, and won’t this undermine whatever subsequent evidence manipulation he manages to perform?

I say no. To explain why, let us consider in further detail the two sorts of practices we have been discussing.

In EVIDENCE AVOIDING, Jones suspects that looking at the handout will give him some bad news, and therefore does not look at it. But at no point need he believe the bad news. He knows that if he avoids looking at the handout with the exam pass rate information, he will be able to persuade himself that his evidence for failure is just too weak. It will not be a matter of 100% conviction, he might have doubts. These patterns of wishful thinking are not only different from Feldman-irrationality; they make sense only because Feldman-irrationality is so
hard to come by. He wishes that he could look at the handout and believe that he will pass nonetheless. Since he cannot, he engages in evidence-avoiding behavior.

Similarly, consider the case of a relationship where there is evidence of infidelity. The partner who fears that the other has been unfaithful might go to significant lengths to remain ignorant of the evidence of infidelity. No doubt this behavior is often motivated by other facts which suggest that infidelity might be taking place, but those might well be insufficient to justify a belief that it is occurring. It is not a simple matter of believing against one’s evidential beliefs—if that were all it amounted to, then why so much effort to remain ignorant?

Here is the general pattern that emerges: We believe that C, and C lends some support to F, which we do not want to believe true. (Perhaps we do not want it to be true, but at the very least we do not want to believe it.) It is not, however, conclusive evidence—not enough to support belief in F on its own—it is rather a mere clue which points in the direction of F. So we are not guilty of Feldman-irrationality in believing C but not F. Fearing that F might be true, after all, is not the same as believing that F is true.

Even so, if the person is aware that this is going on, won’t this undermine his evidence sufficiently? I acknowledge that evidence-avoiding works better the less we attend to how we came by our current evidential beliefs. But while evidence-avoidance’s effect on belief will be more successful the less aware of the avoidance one is, even when one is fully aware of what one is doing, this is not a case of Feldman-irrationality. My evidential beliefs support p. I am strongly committed to p, and it would go badly for me if p turned out to be false. In fact, I fear the possibility that p could be false—even if I believe the probability of its being false is quite low. I fear that consulting a particular source might give me a piece of evidence that undermines

\[58\] This example will be explored further, though in a somewhat different context, in the following chapter.
So I do not consult that source. If I see that source walking down the street, I might cross over to the other side.

Suppose, analogously, that I have a strong fear of defeat in sports. Someone challenges me to a one-on-one game of basketball. I believe that I probably would win, but still decline the challenge nonetheless, because I also believe that I might lose, and cannot tolerate the risk. Similarly, Jones might believe that he will pass, yet still avert his eyes from looking at the data on the pass rate for the exam, because he cannot tolerate the risk of believing that he will fail.

What is going on in WISHFUL THINKING is a little trickier to describe. I said of Jones in that case that he "rethinks things". I will try to describe here what that might amount to. Sometimes, in ordinary rational thinking, after we come to a conclusion, we question the things we believed (our evidential beliefs) which supported that conclusion. This happens in cases where the conclusion seems obviously false. We believe A, B, and C—and together they clearly imply D. But D is absurd. We are sure that something has gone wrong somewhere, but we are not quite sure where.

Consider a case in which three scientific theories which I accept jointly imply a particular experimental result. Suppose that the actual result is wildly different. I repeat the experiment several times and rule out the possibility of experimental error. I conclude that one of the three theories which I accept must be false. I do not instantly disbelieve all three theories—in fact, I may still believe all of them, albeit with a lessened degree of credence, until I examine them more thoroughly to figure out where the problem is.

What we do in these cases is backtrack. Do A, B, and C really support D? If so, we question A, B, and C. Why do we believe them? And we move backward from there. If we were told A, B, or C, we might question the reliability of the source. If we concluded A, B, or C
from other premises, we might move on to examining those. Eventually, we find one to doubt, and thereby eliminate the “epistemic pressure” to believe the absurd conclusion.

Just as we do this legitimately upon arrival at an absurd conclusion, we can also do it illegitimately upon arrival at an undesired conclusion. We submit the relevant beliefs to greater scrutiny. A source whose reliability we have taken for granted in the past is now suspect. We reflect on the ways in which the source could be mistaken, and lose confidence in his judgment. We can also subject ourselves to the same scrutiny—I thought that \( p \) was true, but, perhaps I could be wrong. Perhaps I misread, miscalculated, or thought I saw something I really didn’t see. The only thing that differentiates this process from the acceptable non-self-deceiving one is the motive—whether we are interested in forestalling a merely undesired conclusion as opposed to forestalling an obviously false one.

At the beginning of the wishful thinking process, it is true, the believer believed the undesirable conclusion, and had the evidence to support it. But through this process, he altered his evidential beliefs and thereby undermined his confidence in the undesirable belief. So the objection does not stick for wishful thinking either—it is not a case where our total current set of evidential beliefs fails to support the objectionable belief.

6.4. Objection 4: Given his predicament, Jones ought to believe that he will pass.

Much of what Feldman says in “The Ethics of Belief” suggests that he would offer the following objection: “It is true that Jones in your second and third cases (WISHFUL THINKING and EVIDENCE-AVOIDING) has done what he ought not do. I do not think this is an instance of the epistemic ought, but nonetheless, he has done something wrong and is
irrational in some way. But, once Jones has gotten himself into that predicament, given the
evidential beliefs he has at that time, it seems that the right, the rational, the epistemically proper
thing is to go with the evidential beliefs he has—and believe that he will pass. How can you
claim that he ought not to believe that he will pass?"

A moral analogue may help highlight the force of the objection. Consider a surgeon, who
is scheduled to perform a lifesaving operation on a patient. She is the only one available who is
qualified to carry out the surgical procedure. Three hours beforehand, it is clear that she ought to
perform the surgery at the scheduled time. But suppose she then goes out and gets drunk, and by
the scheduled time is in no shape to be operating on anyone. At that time, it seems, she ought
not to operate on the patient—it would be disastrous if she did.

One might argue that WISHFUL THINKING and EVIDENCE-AVOIDING and similar
cases should be treated in the same way. My past epistemic practices might be lamentable, but
the fact remains that I ought to believe presently what my evidential beliefs support. In the past
it might have been the case that I ought not to believe p, but now that I have self-deceptively
botched my evidence, not believing p would fly in the face of those evidential beliefs, heaping
one sort of irrationality on top of another. If there is no immediate epistemic defect involved,
there is nothing wrong with the believing.

There are (at least) two sorts of reply that could be made here. I will offer the one I like
less first. One could argue that there is an ongoing process in these cases of self-deceived belief.
It is not simply a matter of past history involved. On this view, the self-deceptive processes
described above do not just create the belief and disappear—the work of self-deception continues
afterward to protect and reinforce the belief. I am continually protecting myself from the
evidence out there in the world, blocking my awareness of it somehow. Whatever motivated me
to engage in self-deceptive practices in the first place continues to exert an influence on my thinking. This is easy to see in the example of George Washington discussed above—it seems he did not merely arrive at a self-deceived belief in the past, but rather that he was continually trying to remain blind to the truth, to avert his eyes from the wrongness of slavery, so to speak. So there is an immediate epistemic defect in such cases, although it cannot be simply described in Feldman's terms. It is not simply a matter of failing to believe in accordance with one's evidential beliefs.

I do not favor this response because I do not think we are restricted to making DJB's of the sort I describe only in cases of ongoing tension where the self-deceptive practices play this role of continuing support. Clifford claimed that his shipowner arrives at a "sincere and comfortable conviction," and whatever the other problems with his view, I find his classic example very plausible. I suggested above that we can paint ourselves into an epistemic corner, that we can get ourselves into a peaceful, settled state where our evidential beliefs and the evidence we have ready access to do support whatever it is we self-deceptively got ourselves to believe. I am not sure how common this is, relative to the other scenario of a continuing epistemic struggle, but it seems possible, and the sort of thing that sometimes happens. And it still seems to me in such cases that people are believing what they ought not. But it should be noted that this sort of response could cover many if not all cases of self-deceptive practices and DJB's.

The other reply is to say that I don't think I need to deny the intuition behind this objection. At the very least, as a statement of procedural rationality, it seems there is something correct about it. And I have already acknowledged that we use ought-to-believe statements in various different ways. All I need is that we also use ought-to-believe statements to express
DJB’s like the ones I am describing here, that Feldman’s favored sort of ought-to-believe statement is not the only one and more specifically is not the one behind most of our DJB’s. At the risk of repeating myself, the Feldman understanding of "S ought not to believe p" is unlikely to accommodate most of what I take to be our ordinary DJB’s. First, because it seems to involve a sort of "flatly irrational" defect which seems to make deontological judgments inappropriate, Feldman's weak appeal to role obligations notwithstanding. Second, because the phenomenon seems quite rare, which explains why we engage in self-deceptive practices of the sort I describe. If it were easy to believe something that contradicts your evidential beliefs, we wouldn't need to bother with such elaborate ways of trying to deceive ourselves. Thus, it appears that in many ordinary cases, much like Clifford’s shipowner, we do say of people that they ought not to believe something even when their evidential beliefs support it. Feldman can insist that all such epistemic criticisms are mistaken, but then his account appears unsuccessful as an explanation of our ordinary deontological judgments about belief.

We say of the woman who has joined a UFO cult and is making preparations for the benevolent aliens’ arrival that she ought not to believe that junk. We say of the young man with so-so athletic abilities who plans on being a professional football player that he ought not to believe that career path will work out. We say of the man who has pinned his hopes for his ill brother on a wacky, unproven “miracle cure” that he ought not to believe that it will work. We say of a racist that she ought not to hold her particular racist beliefs. I think we stick by these judgments even after it is pointed out to us that these people’s current evidential beliefs might well support their conclusions. And I think we stick by them even if it is pointed out that in each case the person has arrived at “a sincere and comfortable conviction” and has managed to
insulate himself from the facts so that there need be no continuing self-deceptive process
protecting the desired beliefs.

The force of the analogous moral example of the surgeon can be mitigated by pointing
out that it can speak with two voices. Recall the soccer player from the beginning of section 4. I
said that her teammates can reproach her by saying that "You ought to be playing right now," or
"You ought to have played today." Similarly, we can imagine the surgeon sitting in the hall
halfway through the operation, just beginning to sober up, head in her hands, feeling awful
because she knows she ought to be operating on that patient. We can imagine her stricken by
guilt after the patient dies on the operating table at the hands of someone less qualified, feeling
that she ought to have operated--she ought to have been the one to operate, and she wasn't.

Finally, in response to the concern that genealogy alone should not be enough to earn
negative DJB’s for our beliefs, I would point out that what is doing the work is not merely the
dubious past history of the evidential beliefs, but what that history has done to one’s epistemic
condition. After all, self-deceptive practices attack the connection between the world and one’s
evidential beliefs--the very connection which makes it plausible that we ought to believe what
our evidential beliefs support in the first place. Self-deceptive practices greatly diminish the
epistemic value of one’s evidential beliefs. We might therefore say that my engaging in self-
deceptive practices makes it no longer obvious that I ought to believe what my evidential beliefs
support, at least with regard to the matters I am self-deceiving about.

7. Conclusion

My aim in this chapter has been to explain how our ordinary practice of making
deontological judgments about belief might be justified, despite the dubiously voluntary nature
of belief. I argued that previous attempts to solve the problem by arguing that we have voluntary control over our beliefs or by arguing that control is simply not necessary fell short. Considering the sorts of cases where we are prone to make DJB’s suggested a possible alternative: perhaps what makes us responsible for our beliefs and therefore susceptible to deontological judgments about them is not voluntary control over the beliefs themselves, but rather a broader sort of control over how we conduct epistemic lives, whether we face the facts or retreat from them by way of various self-deceptive practices. (I suggested two possibilities with my examples EVIDENCE-AVOIDING and WISHFUL THINKING—there are undoubtedly many more.) I argued for the superiority of this view over the most fully-worked out alternative view (Feldman’s) and I then addressed various plausible objections to it.

I have contented myself with simply pointing in a general direction in which accounts of DJB’s could go, without offering or defending anything like a thoroughly worked out view of such judgments. To do so is beyond the scope of this chapter, but in closing I will briefly state where I would start in a future investigation of the subject. Very roughly, I would suggest that we might understand “S ought to believe p” along the lines of “There is no ordinary way in which S could fail or cease to believe p without resorting to self-deceptive epistemic practices.” Similarly, I would suggest that we might understand “S ought not to believe p” along the lines of “There is no ordinary way in which S could continue or come to believe p without resorting to self-deceptive epistemic practices.” Of course, the devil is in the details, but something like this appears to be behind our everyday deontological judgments about belief. When we make DJB’s about others we are assessing whether they could hold or avoid holding a particular belief “innocently,” or whether they would need to resort to self-deceptive practices to do so.


Belief and Commitment

1. Introduction

Good believing, by and large, is believing that is supported by the evidence. How exactly to spell that out is a matter of controversy I will not address here. But I take it to be uncontroversial that insofar as someone's beliefs are proportioned to and responsive to the evidence she has available to her, something is going right with her beliefs epistemically speaking. Someone whose beliefs are unresponsive to the facts that he is presented with, has beliefs which are epistemically defective in virtue of that fact.

There are things we do which are likely to influence the quality of our beliefs, and we might say that the epistemic norms tell us what to do in order to have beliefs of the good kind. As Simon Keller puts it, "An epistemic norm is a truth about what you should do, insofar as you are an agent who forms and holds beliefs, and who is answerable to reasons for forming and holding beliefs." Such a norm might stipulate how I should handle the evidence I am presented with. Do I try to avoid looking at it? Do I try to ignore it, explain it away? Do I overscrutinize it skeptically, and doubt it unreasonably? Do I underscrutinize it, accepting it over-eagerly? All of these practices, I take it, are a form of epistemic defect, a violation of an epistemic norm. The self-deceiver believes what his evidential beliefs support, but he has managed to massage and manipulate his evidential beliefs to lead to his desired conclusion, which prevents his beliefs from being sensitive to the truth. He violates a norm which directs us to take evidence as it

---

comes and follow where it leads in forming our beliefs, rather than manipulating it through mental gymnastics in order to arrive at a desired belief.

One might object to calling this norm epistemic, as it does not specifically address our believing, but rather how we direct and influence the process of forming beliefs. I am inclined to consider it epistemic as it deals with practices which cause one's beliefs to fail to be responsive to the truth and make justified belief all but impossible. However, nothing of significance in what follows hangs on the label, and I do not intend to argue for it here.60

Conducting our epistemic lives so that our beliefs conform to the evidence we are presented with appears to clash with our ability to pursue, enjoy, and honor certain things that we value, such as social relationships, loyalties to individuals and groups, and ideals that we hold. In everyday life, we often find irrational belief cropping up in these areas—about what is a person more likely to deceive herself than her friends, her family, the groups of which she is an affectionately devoted member, and her most cherished hopes and ideals? Yet we look on such irrationality quite indulgently, even approvingly. There is nothing wrong, it seems, with someone who has a higher opinion of his friend than the friend objectively merits—in fact, it is plausible that there is something wrong with a friend who lacks that bias.

Taking off from these intuitions, it has been argued that the perfectly rational believer, one whose beliefs never deviate from what the evidence he is presented with suggests is most likely to be true, must miss out on important goods essential to a good human life. Some have pointed out the conflict and left it at that, others have asserted that the other norms and values trump the epistemic ones, and argue in favor of irrational, counter-evidential believing.61

61 The former include Keller and Stroud; the latter include James and Meiland. No one to my knowledge has argued that our social relationships and loyalties must be sacrificed.
While granting that self-deceptive beliefs are prevalent in these areas, I argue that they are not necessary in order to secure these goods. Where it is suggested that we need to self-deceptively or irrationally believe a proposition in order to be a good friend, citizen, etc., there is another attitude available to us, which I will call propositional commitment, which allows us to obtain these goods and discharge our obligations to the individuals and groups to whom we owe them without the costs of self-deception and believing against the evidence.

It might be thought at the outset that this is a gratuitous complication. What is so bad anyway about self-deceptive / counter-evidential beliefs in these harmless sorts of cases? As long as we reject W.K. Clifford’s overly earnest conviction that “it is wrong always, everywhere, and for anyone, to believe anything upon insufficient evidence,” and that any goods we enjoyed as a result of such irrational believing would be pleasures “stolen in defiance of our duty to mankind,” who cares? So before I set out and defend my alternative, I will show that the difficulties and costs involved in simply endorsing irrational and self-deceptive belief in these cases are significant enough to justify proposing an alternative. But first I will describe the argued-for conflict in more detail, so that we have a better grip on what it is.

2. The Problem

---

63 Ibid., 344.
It has been argued that sometimes we ought to believe irrationally, that we ought to believe propositions which we have no epistemic justification for believing, no (or insufficient) evidence to support them.

Advocates of such a view are generally not trying to argue for a different set of epistemic norms. Rather, their position is analogous to Susan Wolf's in "Moral Saints," where she argues that while morality may be very demanding, we should only be so moral.64 There are, I think, reasons to worry about "moderation in being moral," but these seem unique to morality. Moral reasons have a weightiness and decisiveness that other kinds of reasons lack, so the analogous epistemic position need not face those problems. We might say that certain goods are essential to a good human life, and if epistemic norms stand in the way of these, so much the worse for epistemic norms. There may be other reasons to worry about this view, but at the outset, "So much the worse for the epistemic norms!" is nowhere near as jarring as "So much the worse for morality!"

Instances of apparent conflict are legion. The best-known and most thoroughly discussed cases involve conflict between epistemic norms and the norms and values connected to social relationships. Two kinds of argument are made here. First, it is argued that the goods of social relationship are available to us only if we fail to conform our beliefs to the evidence to some degree. Initiating a relationship, William James argued, requires a degree of trust that goes beyond what the evidence justifies.65 Even after the relationship has been initiated, it is plausible that an epistemically unjustified bias in our beliefs about those we are in relationship with might be required in order to fully enjoy the goods of that relationship. It seems that relationships—

---

whether between friends, lovers, or family members—cannot be fully enjoyed as long as the parties to the relationship insist upon complete epistemic integrity, a strict adherence to what is justified by the evidence. That would be to treat one’s friends like a stranger, like “just anybody.”

In contrast to that consequentialist type of concern with what having beliefs strictly conformed to the evidence might cost us in terms of relationships, there is a deontological concern as well. It could be argued that we owe an epistemically impermissible partiality to our friends, as one of the demands of loyalty. Regardless of whether believing \( p \) (an unpleasant proposition involving my friend) would hurt our relationship and deprive me of its pleasures, it might be that I ought not to believe it simply because we are talking about my friend. It is not simply that I would be a foolish friend to believe it, on this view—I would be a bad friend.

Wolf’s claims about the unappealingness of moral saints are easily adapted to these kinds of cases. The “epistemic saint” who responds exclusively to the evidence, the person who is entirely unresponsive to the dynamics of relationships in his believings, the person who knows nothing of trust or “faith” in another person, is starkly unattractive to us. The person who responds impartially to evidence that someone dear to him has done something horribly wrong is a ghastly figure to us. We demand, it seems, some degree of denial or resistance to the evidence in such cases as part of being human. If we concernedly query a friend about some rumor we have heard about him, he will likely reply indignantly, “How could you even think such a thing about me?” He may not be so much accusing us of a lack of evidence as he is accusing us of being susceptible to the evidence in a situation where we ought not to be, or at least more susceptible than we should be.
Although the most plausible examples of conflict involve human relationships and social life in these sorts of ways, there may be others as well. The partiality and reduced sensitivity to evidence apparently demanded by relationship loyalties might be demanded by other loyalties as well—loyalties to a group which one is a member of, or one’s country, for example.66

Perhaps holding certain beliefs related to one’s ideals independently of their justification by the evidence is to be valued as well. For example, we might admire someone for their belief in the fundamental decency of human beings while being far from confident that such a belief is justified. Other sorts of socially hopeful beliefs and big-picture optimism might qualify as well. We might think it laudable (or at least permissible) to believe with Martin Luther King, Jr. that “the arc of the moral universe is long but it bends toward justice.” Of course, we do not always respond to optimism favorably—we are as likely to scorn people for being Pollyannas or ostriches with their heads stuck in the sand as we are to admire them for their “faith” in humanity or democracy or non-violence or the progress of justice. Nonetheless, we sometimes treat idealistic and optimistic beliefs (when they are not immediately self-serving) as praiseworthy, sometimes even especially praiseworthy because of their flying in the face of the evidence.

In short, although the kinds of cases may vary in their persuasiveness, there appears to be a conflict here. It seems that there is something appealing or good or right (I do not think our intuitions are sharp here, so I will avoid specificity) about violating the epistemic norm which tells us that our beliefs ought to be conformed to the evidence in these sorts of cases, and it seems that there is something unappealing or bad or wrong about adhering to the norm.

66 But see Simon Keller’s worries about this in his “Patriotism and Bad Faith,” *Ethics* 115 (2005), 563-592.
3. What Do Cases of Conflict with Social Relationships Look Like?

Although I find the ideals-related cases mentioned above intriguing, I will focus primarily on the social relationship cases, as they seem to be the most plausible instances of acceptable deviation from evidentially-supported belief. In this section, I will briefly describe the sorts of cases found in William James’s classic lecture “The Will to Believe,” as well as in more recent contributions by Jack Meiland and Simon Keller. The range of examples will help us understand the conflict and provide fodder for discussion in following sections.

While James’s main concern in his lecture is religious belief, he devotes attention to social relationship cases in passing, looking at self-fulfilling nature of apparently unjustified beliefs at the outset of relationships. He claims that when we are willing to believe that others like us, though on insufficient evidence, they are more likely to like us. “The previous faith on my part in your liking’s existence is in such cases what makes your liking come.” 67 I may know nothing about you, yet if I believe that you like me or are at least likely to respond to me favorably, then I will find it much easier to be friendly and naturally warm with you, which will make me more attractive to you, and make you more likely to respond favorably. 68 In contrast, “a man who in a company of gentlemen made no advances, asked a warrant for every concession, and believed no one’s word without proof, would cut himself off by such churlishness from all the social rewards that a more trusting spirit would earn. . . .” 69 On this view, to be epistemically rigorous in our beliefs about others and our relationships with them,

67 “The Will to Believe,” 730.
68 There are, of course, limits. If my unjustified confidence in the likelihood that others will respond favorably to me leads me to be pushy and overfamiliar, I am more likely to produce an annoyed reaction than a friendly one. And if the person in question is actively but politely trying to signal their dislike or lack of interest, my obliviousness to that is similarly unlikely to go over well.
69 Ibid., 733.
accepting only what the evidence warrants, would bar us from the pleasures of social interaction by preventing us from having any social relationships to begin with.

Jack Meiland’s two main examples in his “What Ought We to Believe?” focus instead on maintaining a valued relationship in the wake of behavior which threatens the relationship. The first gives us Jones and Smith, business partners and close friends. Jones “comes into the possession of evidence which is sufficient (in anyone’s eyes) to justify the belief that Smith has been secretly siphoning off money from the business.” Jones knows he cannot conceal his feelings from Smith, and therefore if he believes that Smith stole the money, the relationship and the partnership will end. Valuing the relationship, he concludes that he ought to believe that Smith did not steal the money, and therefore, according to Meiland, he does so believe. Meiland’s second case is “the classical case in which a wife finds a blonde hair on her husband’s coat, a handkerchief with lipstick on it in his pocket, a scrap of paper with a phone number scrawled on it, and so on until everyone would agree that the evidence is sufficient that the husband has been seeing another woman.” The wife believes that this is just a rough patch, and if they can get through this, their marriage will “weather this storm.” But, like Jones, the only way she can keep up her end of the relationship is to come to believe that her husband is not cheating—she cannot hide her feelings. And so, Meiland claims, she decides to believe that he is not cheating, and she decides rightly. He thinks it is clear that both of these people ought to believe against their evidence. Specifically, because of their valuing of their respective relationships or the loyalties due to their respective partners, they ought to act in a way that

71 Meiland embraces a strongly voluntaristic conception of belief, which I do not share; I will discuss this further in section 5.
72 “What Ought We To Believe,” 16.
allows the relationship to continue, and if a particular belief is necessary for that, then they ought to hold that belief.\textsuperscript{73}

Simon Keller's examples, in his "Friendship and Belief," are more subtle than Meiland's. His first is borrowed from the sitcom \textit{Friends}:

Joey and Chandler are driving from New York to Las Vegas, where Joey has landed an acting job that he hopes will be his big break into show business. They are playing a game, which involves Joey asking questions and Chandler giving immediate, unreflective replies...One of Joey's questions is, "Is this job going to be my big break?", and Chandler, before he can catch himself, answers "No." A crisis in the friendship ensues; Joey feels betrayed, and Chandler feels like a betrayer. Joey expels the remorseful Chandler from the car, and drives to Vegas alone.\textsuperscript{74}

Keller argues that we can make sense of the two friends' responses by saying that because of their friendship, Chandler ought to have believed that the job was going to be Joey's big break.

His second example involves two friends, Rebecca and Eric.\textsuperscript{75} Rebecca is scheduled to give a poetry reading at a café and is nervous about it. She hopes to capture the attention of a literary agent who will be there. She asks Eric be in the audience, presumably to supply a friendly face, a sympathetic ear, and moral support. Eric knows nothing about the quality of Rebecca's poetry, but he generally has a rather low opinion of the poetry read at that venue. Yet when he goes for Rebecca's sake, Keller argues, he ought not to believe the poetry she is going to read will be bad, and furthermore, that his friendship should have a subtle way of influencing his beliefs so they are more favorable than they would have been otherwise. Friendship demands a partiality, a presumption of favorability towards one's friend, that requires one to believe against the evidence, to put oneself in situations and frames of mind which one knows will lead to false beliefs.

\textsuperscript{73} Ib., 20
\textsuperscript{74} "Friendship and Belief," 329.
\textsuperscript{75} Ib., 332-3
Having considered some examples of the conflict, where it seems we are required to deviate from good, rational believing in order to obtain relationship goods and uphold our loyalties, I will examine two kinds of responses to the conflict.

4. Feldman: The Two Kinds of Values are Incommensurable

Richard Feldman implicitly acknowledges the conflict when he claims that talk about what we ought to believe must be disambiguated in terms of moral, epistemic, and prudential oughts, as there is no all-things-considered sense of the term “ought.” He dissolves the conflict by saying only that we epistemically ought to believe what is supported by our evidence, leaving it open that we might morally or prudentially ought to believe otherwise, and insists that weighing the different considerations would be impossible. 76

This approach is unsatisfying enough that it should only be accepted as a last resort. As rational, responsible, reflective beings, we want to have a coherent and reasonable plan for negotiating conflicts between kinds of values and norms when they clash in our lives. We do not want to shrug our shoulders indifferently and flip a coin every time we have to choose between them.

After all, we know there are better and worse ways to rank different values and sets of normative claims, in terms of deciding which trump which. If we find ourselves in a situation where the norms of good chess-playing, for example, tell us to do something that conflicts with what serious moral norms tell us to do, it is obvious that we must go with the moral norms. Likewise, violent cruelty to other human beings simply could not be justified on the grounds that

it was necessary to create an especially great work of art. We would not just say, “Well, morally it is appalling, but artistically it’s quite grand,” as if that were all that could be said. Similarly, although perhaps not so strikingly, when it comes to the conflict between epistemic values and other values, some ways of dividing the spoils between them are simply wrong, full-stop. Someone who always believed whatever he found most pleasing to believe would not be merely epistemically vicious. But if we are not going to take the escape route of declaring epistemic norms and values to be incommensurable with other norms and values, then something more must be said about cases where they appear to conflict.

5. James and Meiland: We Are Free to Believe Against Our Evidence in These Cases of a Clash, and in Fact, Ought to Do So.

James and Meiland offer a different solution. They claim that where epistemic norms interfere with our ability to pursue and honor our social relationships, so much the worse for epistemic norms. They claim that we have voluntary control over at least some of our beliefs, and that it is appropriate to exercise such control to believe against our evidence or independently of it. James famously suggested that in cases where the evidence was inconclusive, one was free to bring in considerations other than evidence (in particular, one’s “passional nature”) in willing what to believe. Meiland insists that considerations beyond evidence should enter into every “decision” to believe.

For various reasons, these recommendations will not work. First, as I will argue, their voluntaristic notion of belief is in need of a far stronger defense than they supply. Secondly,

77 “The Will to Believe,” 723.
78 “What Ought We To Believe,” 21.
even if the position is modified to avoid that problem, those who hold it face a dilemma where both alternatives are extremely costly and unattractive. Showing that the view has unappealing consequences is not sufficient to refute it outright, but it should justify an attempt to offer an alternative account of what is going on in these cases.

Objection #1: They fail to explain how belief can be voluntary in the ways they suggest.

Both James’s and Meiland’s accounts assume voluntary control over one’s beliefs. James speaks of willing to believe; Meiland speaks of deciding to believe. Yet neither attempts to substantively address the controversy about how such control over one’s beliefs could be possible. We might initially charitably assume that James was simply unaware of the controversy, yet he seems perfectly aware of the problem when it comes to beliefs for which we have compelling evidence:

> Does it not seem preposterous on the very face of it to talk of our opinions being modifiable at will? Can our will either help or hinder our intellect in its perception of truth? Can we, by just willing it, believe that Abraham Lincoln’s existence is a myth, and the portraits of him in McClure’s Magazine are all of some one else? Can we, by any effort of our will, or by any strength of wish that it were true, believe ourselves well and about when we are roaring with rheumatism in bed, or feel certain that the sum of the two one-dollar bills in our pocket must be a hundred dollars? We can say any of these things, but we are absolutely impotent to believe them; and of just such things is the whole fabric of the truths that we do believe in made up—matters of fact, immediate or remote, as Hume said, and relations between ideas, which are either there or not there for us if we see them so, and which if not there cannot be put there by any action of our own. 79

What is odd about James’s position is that he thinks that cases in which we have compelling evidence are somehow special. He acknowledges that if we are struck by overwhelming

79 “The Will to Believe,” 719.
evidence for $p$ and grant a high degree of credence to $p$, that we cannot will to believe not-$p$. Yet it seems more plausible that the problem is a general one, applying to all degrees of compellingness of evidence and all degrees of credence. We cannot straightforwardly adopt a degree of belief stronger than what we perceive the evidence to justify. We must trick ourselves somehow into thinking that our evidence is better than it is, putting ourselves into situations that will control the evidence we encounter, etc., in order to modify our beliefs.

Meiland also mentions the controversy over whether or not we have control over our beliefs in passing, but argues that it is as much a problem for his evidentialist opponent (who will argue against him that what we ought to believe is determined by the evidence) as it is for him, as ought implies control of some sort. He thus considers himself freed from any responsibility to show how it could work. However, if I am right that we can justify ought-to-believe judgments without having direct control that allows us to “decide to believe” in the face of the evidence, it appears he does owe us some sort of explanation.

His next step is almost identical to James’s, saying that in some cases of belief we have conclusive evidence, which forces or compels belief, but in some cases we merely have sufficient evidence, which justifies belief but does not compel it. When we have only sufficient evidence, we can decide whether or not to believe the proposition that the evidence supports. He offers the example of a defense attorney in a criminal trial who believes his client to be innocent in the face of sufficient (but not conclusive) evidence to the contrary, and concludes that “Sufficient evidence does not result in some logical or psychological necessity which forces the belief in question upon the person.”

---

80 “What Ought We To Believe,” 16.
81 See Chapter 2, “Deontological Judgments about Belief” 49-53.
82 “What Ought We To Believe,” 17.
Yet the details of the example raise more questions than they answer. He claims that this defense attorney believes his client to be innocent while believing “the jury’s verdict to be fair and eminently justified,” and believing that the evidence is “sufficient to justify a certain belief (that his client is guilty).” In other words, according to Meiland, his attorney believes his client’s guilt has been proven *beyond a reasonable doubt* and yet believes him to be innocent. This seems rather unlikely, but even if the attorney could do this, it has little to do with willing to believe. Perhaps he believes that he knows something about his client that the jury did not or could not know—perhaps because of some evidence which was inadmissible in court, or perhaps simply because of how well he has gotten to know his client. But then his belief differs from the jury’s verdict because he thinks he has evidence they do not. The evidence they had justified their belief and verdict, but in light of what he knows, he (justifiably) believes otherwise. Or perhaps over the course of his career, he has gradually deformed himself epistemically so that he cannot respond rationally to the evidence against his clients. But neither of these explanations suggests a voluntaristic flexibility in the attorney’s beliefs, and it is not clear what other interpretations could make Meiland’s claim intelligible.

James and Meiland thus paint a picture where we can believe what we choose except when a burden of overwhelmingly compelling evidence stands in our way. Belief is voluntary except when it is constrained by especially good evidence. Yet while each seems aware that there are apparent problems lurking in the neighborhood, neither takes up the challenge of explaining why we should think such a thing to be possible.

It is very hard to see how it could be possible. After all, there are many propositions for which I do not even have sufficient evidence for or against. Yet none of these seem any more capable of being voluntarily believed or disbelieved than any proposition for which I have
conclusive evidence. Take, for example, a proposition about the size of the population of Luxembourg (rounded to the nearest hundred, say.) I have no sufficient evidence to believe any such proposition, but that lack of evidence does not allow me to believe anything. The problem is not simply a lack of reason or motive—a villain could threaten to kill or torture someone unless I believe a particular proposition about the population of Luxembourg, but I still could not believe it. If this is not possible, how can choosing to believe the negation of a proposition for which we have sufficient evidence be possible? Believing against one’s evidence—as Meiland insists is possible and commends—could only be more difficult than believing where one lacks evidence entirely. Perhaps a story can be told about the cases which concern James and Meiland, but in the absence of such a story their claims are hard to accept.

Our standards for what counts as conclusive evidence may depend on the context, but we cannot choose which standards we are going to apply. Perceptions of a given context may differ, and people with different epistemic temperaments may apply different standards, but given my perception of the situation and my epistemic temperament, my standards for evidence are beyond my immediate control. I can influence my standards and epistemic temperament in the long run, but the effects will be extremely broad, and that sort of control can hardly be described as deciding to believe anything.

Having addressed this controversy over how voluntary belief can be in the previous chapter, I will not discuss it in more detail here. I contend there that we have responsibility for our beliefs (and can be subjected to deontological judgments regarding our beliefs) only in virtue of an indirect ability to influence what we come to believe. Specifically, we can exercise control over our beliefs by engaging in self-deceptive practices that manipulate our perception of the evidence, and cause us to think that the evidence supports a conclusion other than the one it
actually does. Both James and Meiland depend crucially, I think, on belief being voluntary in a way that it is not.

Objection #2: Believing against one’s evidence must either be conscious or subconscious, and affirming either is quite problematic and costly.

Supporters of the James/Meiland view could respond by modifying the view so that it is instead an endorsement of self-deceptive practices in the sorts of cases where James and Meiland endorsed willing/deciding to believe in the absence of evidence or against the evidence. “No problem,” they might say, in response to the above objection. “We can easily put our claims in terms of the acceptability of using self-deceptive practices to produce certain beliefs. It is permissible to engage in self-deceptive practices (such as avoiding the evidence and wishful thinking) in these sorts of social relationship cases.” But this approach will not work either. In particular, anyone who recommends self-deceptive practices is caught on the horns of a dilemma. Either what they have in mind is a conscious and circumspect sort of self-deception, or a stealthily subconscious and hidden sort of self-deception, and both of these are quite unattractive.

Consider the first possibility—that they are recommending a conscious, circumspect, and thoughtful decision to engage in self-deceptive practices. The problem is that believing things (or even considering whether to believe things) for reasons unrelated to their likely truthfulness or falsity must be parasitic on a general practice of believing justified by the evidence. It is not clear how aware of this James is, but Meiland denies it outright when he suggests that the implications of our beliefs for our personal relationships and other things that we care about

---

83 Incidentally, the objections in this section would apply equally to the sort of voluntary believing against the evidence dismissed in the previous section, as well as bringing about beliefs through self deceptive practices.
should enter into every decision about what to believe. ("So we may put the matter in this way: extra-factual consideration should be consulted in every case to see which factor should determine belief; and in some of these cases extra-factual considerations will themselves be among the factors that should influence belief.")

I do not keep detailed mental records of how I came by my beliefs. In trusting my beliefs, in using them to decide how to act, I assume that I came to hold them because I had good evidence. If I had a habit or a conscious practice of coming to believe things for reasons that are independent of the evidence of the likelihood of their being true, then in the case of every belief whose origin I cannot remember I would have to ask myself, "Did I come to believe this because I thought it likely to be true, or for some other reason?" We might occasionally succumb to self-deceptive practices out of epistemic akrasia, but the idea that we could embrace them consciously and wholeheartedly is problematic. A hard-line consequentialist might insist that one ought to hold whichever belief would have the best consequences, but no one could consciously take this as a general rule for themselves. (This seems even more evident than the claim that no one could take consequentialism as a general rule for deciding what to do.) Disavowing self-deceptive practices is what makes them successful; I believe my beliefs, trust them, am influenced by them, and act on them because I think I am not a self-deceiver, not the sort of person who believes against the evidence. In order to successfully deceive myself, I need to believe that the self-deception is the sort of thing I don't do.

This much is hardly surprising. But the alternative, the claim that while we cannot acknowledge or be aware of our self-deception it is still something we sometimes ought to do,

---

84 "What Ought We To Believe," 22.
85 The conscious employment of self-deceptive practices thus suffers the same sort of difficulty that "willing to believe" faces. See Bernard Williams, "Deciding to Believe," in Problems of the Self (Cambridge: Cambridge University Press, 1976).
faces its own difficulty. Specifically, the epistemic contortion of self-deceptive practices would do undesirable things in the long run to our epistemic and rational well-being. Perhaps we could justify particular exceptions for important relationships, matters, and events, but here we are assuming that we are not conscious of when and where we are allowing ourselves to self-deceive. Given that self-deceptive practices must operate “under the radar,” we cannot regulate or be aware of when we are or are not using them.

Engaging in self-deception involves skills or knacks that one grows more proficient in with time, and ruts that deepen and become easier to fall into with each repetition. Perhaps I start self-deceiving only when the costs of being open to the evidence are things like emotional devastation, or the ruin of relationships. But over time, my standards get lower and lower. Eventually the mere inconvenience of a belief might be a reason not to hold it, or the fact that a belief would make me happy would be a reason for holding it. One therefore cannot protest that they are only recommending self-deception in certain cases. If we could confine our self-deceivings neatly, they would not succeed; if they are to be successful, we cannot be aware of (much less control) what sorts of circumstances we are using them in, and thus are inviting them to spread in our epistemic practices to a degree that James and Meiland should find worrisome. A capacity for self-deception is a bad thing to develop, simply because it cannot be directed solely toward the cases where it might be justifiable to do so.

6. Belief and Commitment

Self-deception thus turns out to be extremely problematic even in these social relationship cases. There are reasons for concern which go beyond Clifford’s perhaps excessive
scruples about the epistemic. Even if we acknowledge the importance of other kinds of norms and values—in particular the values of social relationships—endorsing believing against one’s evidence (either directly or indirectly through self-deceptive practices) comes at a very high cost. This is because of the nature of belief—if we could simply and self-awarely believe against the evidence, it might be acceptable to do so, but we cannot. We must engage in self-deceptive practices, which insofar as we are conscious of them will undermine the practical connection between belief and action, and insofar as we are unconscious of them will creep into other areas of our epistemic lives. Counter-evidential believing, while perhaps not obviously wrong in and of itself, is a troubling price to pay.

What I will argue in the rest of this chapter that we do not need to pay that price. We do not need self-deception or counter-evidential believing in order to secure the relevant goods of social relationship and maintain the relevant loyalties. What we do need is something I will call propositional commitment, by which I mean trying to have the attitude (toward a person, group, institution, or perhaps even an ideal) one would have if one fully believed a particular proposition, while aware of the inability of one’s evidence to support believing that proposition, and therefore not believing it. While we ought to believe in accordance with our evidence, our commitments are not so constrained, and we are free to commit ourselves under certain conditions to propositions for which we do not have conclusive evidence, and perhaps even to some for which we have evidence which speaks against them. We can portray the sorts of deviations from epistemic purity that James, Meiland, and Keller discuss as examples of commitment, not of belief—a modification of their views which captures the best elements of them while avoiding the problems described above.
What does it mean to try to have the attitude toward a person that one would have if one believed a particular proposition? It involves action, thought, and feeling. We try to act towards the person (and perhaps toward others as well) as though we believed the proposition. We try to conduct our mental lives (insofar as they are under our voluntary control) as though we believed the proposition. And we try to feel towards the person as though we believed the proposition. Thus, if I am committed to the proposition that my friend is honest, I will try to treat her as I would if I believed she were honest and speak of her to others as if I believed she were honest. I will try to think of her as I would if I believed she were honest—perhaps dwelling on the evidence in favor of her honesty, while trying not to call to mind information that calls her honesty into question. And I will try to even feel toward her as I would if I believed she were honest. While this is not the same as believing her to be honest, I will argue that it may be nearly as beneficial in allowing me to preserve the relationship I care about. Unlike the dubious capacity for voluntary belief, we know that such a thing is possible; unlike the dangerous capacity for self-deception, it can be done transparently and self-awarely.

I would emphasize that commitment is a matter of trying to have the attitude that we would if we fully believed the proposition, not simply a matter of having it. It is a matter of effort, and it seems likely that we are rarely if ever completely successful—the stronger the evidence is against the proposition, the harder time we will have, no doubt. There is usually a nagging tension with the facts as we are aware of them, a sense of struggle.

Some might find my terminology objectionable. Commitment, they might say, is something we have to people, institutions, and ideals—not to propositions. I happily acknowledge that commitment is primarily directed toward those things. Yet I would suggest that we can have a commitment to propositions in virtue of the commitments we have in the
primary sense. Because of my personal commitment to my friend, I might be committed to the proposition that she is honest, even when the evidence casts doubt on it. I thus see propositional commitment as a derivative of commitment in the ordinary sense. It seems to me useful to be able to talk about it as commitment to propositions, as it involves acting, thinking, and perhaps feeling as though something were the case.

7. What Does Commitment Look Like?

Having very roughly sketched above what propositional commitment is, I will here describe some cases of it in a way that will hopefully fill in the sketch.

7.1 Commitment and the Beginning of Friendships

As we saw above, James argued that we needed to will to believe beyond what is justified by the evidence in order to start new friendships, romances, etc. But I think commitment would suffice here. Consider shy Michelle, who wishes to make Sarah’s acquaintance with an eye to possible future friendship. She tries to have the attitude toward Sarah that she would if she fully believed Sarah were friendly. She approaches and introduces herself to Sarah in a friendly (but not over-familiar) manner, with an openness that carries with it the possibility of being shot down or disappointed to some degree. This “going out on a limb” is self-aware acting beyond what the evidence justifies—it reflects a commitment to the proposition that Sarah will respond favorably. In order to have a chance of success, of course, this act must be accompanied by support from her thoughts. Unless she is a splendid actress, her friendly overtures will not have
their desired effect or come across as very friendly at all if she dwells on the risk she is taking. So she focuses on imagining and anticipating a friendly reaction, and drives the other possibilities from her mind. She turns her attention to cues which suggest a favorable response (perhaps Sarah is holding a book that Michelle has recently enjoyed, or perhaps her shoes reveal a similar taste in footwear) while attending less to cues which might point towards an unfavorable response (such as Sarah’s frown as she thumbs through the book in question). This might fall well short of belief, but it is often enough to enable a successful performance which gets the friendship off the ground.

This is not, of course, a phenomenon which occurs only at the inception of friendships. It seems to be an instance of a kind of self-fulfilling confidence, which occurs in a wide range of interpersonal interactions as well as elsewhere. A teacher, for example, will teach better if he is committed to the proposition that his class thinks he is interesting, insightful, a good teacher, and so on. But I would suggest that despite the diversity in those cases, they can all be understood as requiring that we try to have the attitude toward someone or something that we would if we believed something. Irrational belief is not required to start a relationship or to prepare oneself for any other sort of performance or test where confidence is useful.

7.2 Commitment and Dealing with Evidence or Reports Reflecting Poorly on a Friend

Next, recall how Meiland argued that the wife and the business partner must believe their respective partners to be innocent in order to maintain the relationship. The psychologies of the characters in his examples are of course his to stipulate, but it seems that commitment, trying to have the attitude that we would have if we believed them to be innocent, could play the necessary role for most of us. Consider Mary, who is Joe’s friend. Word has been going
around that Joe is in some sort of trouble. Mary wants to keep the relationship, and she believes that Joe is basically a good guy, and whatever is going on, if anything, cannot be too serious and must be a mere blip in his character and his life. Mary wants to be a good friend to him in the meantime, and so she acts as though nothing is wrong. (We might assume that she has good reason to believe that confronting Joe or asking him about the matter would only hurt the relationship, whether because of her reaction or his.\textsuperscript{86} ) She sensitively avoids certain subjects with Joe, and even keeps her distance from friends who might give her more “dirt.” If in their presence, she tries to defend him by offering alternative explanations of what has been going on. At the same time, she controls her thoughts about the matter to some degree. She deliberately steers her mind away from speculations about Joe’s guilt and what he might have done. Whatever information she does come across, she tries to interpret as positively as possible. She continues to try to have the attitude toward him that she would have if she were sure that he wasn’t involved in some sort of shady dealings, if she were completely certain of his upstandingness, but, I think, in a way that falls short of belief. She focuses on fond memories of old times, and Joe’s good qualities, reminding herself that he (they) will get through this somehow. And, in fact, they do. All of this is possible, I would argue, without her believing that nothing is wrong.

Note that this is not simply a matter of Mary’s finding the rumors or allegations hard to believe in virtue of what she knows about Joe, in virtue of the other evidence she has. She might not have that long a history with him. She might be well aware of the tendency of human beings

\textsuperscript{86} Meiland’s cases suggest that the relationship problem would be with the person doing the committing—they could not continue the relationship if they believed thus-and-so. But I think the breakdown of relationship might well involve both parties. The defensiveness of the “guilty” part might play as big a role as the hostile suspicions of the “innocent” party. We needn’t assume that Mary is engaged in commitment, therefore, because she believes she can’t handle the bad news. That might be the case, but it might also be that Joe couldn’t handle her believing the bad news—he might care a great deal about how she thinks of him, how she looks at him.
to have dark sides and secret lives. She might even be aware that she has reason to suspect that something is wrong. We can hold this sort of attitude toward our friends even when we suspect that whatever they are being accused of might well be true of them, when we know full well that they are capable of whatever it is.\textsuperscript{87}

While Meiland’s examples (and mine which attempts to mimic his in certain respects) focus on moral failings in one of the people in the relationship, I would point out that this sort of case could involve other matters as well. Any fact about Joe which might make Mary uncomfortable around him, or which might make Joe uncomfortable around her if he thought she were aware of it, could play the same role and motivate the same sort of commitment.

Note also that it is not enough for Mary to simply repress and block negative thoughts and suspicions about Joe. She must also work toward looking at him in a positive light, trying to have the attitude she would have if she believed everything was all right. The same would go for the wife in Meiland’s infidelity case, if she were to try commitment instead of self-deception—it would not be enough for her to simply try to not think about the evidence and think of other things instead, she must try to have the attitude she would have if she believed that her husband were faithful. Merely trying to turn one’s mind away from the disturbing facts would not be enough to keep the relationship going. This is because relationships require trust—neutrality is not enough.

7.3 Commitment and How We View Our Friends’ Prospects and Performances

\textsuperscript{87} Sarah Stroud in her “Epistemic Partiality and Friendship,” \textit{Ethics} 116 (2006), 498-524, points out that the literature sometimes distorts matters here by focusing on extreme cases, like those in which the friend has been accused of murder. We are reluctant, she notes, to accept a bad report about a friend’s character or behavior even when it involves “more everyday sins,” such as sleeping with someone and then not returning their phone calls.
Finally, remember how Keller argued that the norms of friendship require us to have irrationally biased beliefs about our friends. According to him, Chandler ought to believe that Joey’s job will be his big break into a successful acting career, and Eric ought to believe more optimistic and favorable things about Rebecca’s poetry reading than the evidence justifies. Again, I think commitment is more than able to explain what goes on in the mind of a good friend.

Consider John, Susan’s friend, who goes to her cello recital and listens sympathetically. He tries to have the attitude toward her and her performance that he would have if he fully believed it were an excellent performance. This takes little in the way of action, but more in the way of thought. Mistakes or flaws in the performance that might otherwise turn him off from further interested listening he tries to dismiss from his mind as soon as he notices them. In contrast, he consciously dwells on particularly successful passages. He reads all of her interpretive decisions as hopefully and charitably as possible. As a result, when the performance is over, he has a far more favorable impression of it than he otherwise might. Yet at the same time, because he is conscious to some degree of what he has been doing in listening, of how the friendship has influenced his listening, his impressions do not really have the status of belief. If asked for his assessment of the performance, he will share his impressions, but may well qualify them with the acknowledgment that he is, after all, her good friend—especially if it clear that an impartial, objective assessment is what is being asked for.

It is not simply a matter of finding things to like about a person or performance. Parents might watch their young daughter’s (awful) ballet performance and think "Oh look at how she's enjoying herself. Look at her enthusiasm and vigor!" and so forth. The qualities of the performance are simply ignored. They do not try to have the attitude that they would if things
were better than they are, if their child had more talent than she does. But surely this is not what Rebecca is looking for from Eric, or Susan from John. Our friends would feel insulted if we tried to commend them or defend their performances to others in those terms. That is at the heart of what is relevant here—what the person thinks of his own abilities, chances, or performance. If my friend happily knows herself to be a poor excuse for a poet, actress, or musician, and performs simply to amuse herself, I need not commit myself in any way. But if she longs to be successful, hopes to do well, and wants to impress, as a good friend my attitude should be influenced by that. Our friends want our attitudes to mirror their own, as I will discuss further below.

What I hope these descriptions show is that in a variety of instances, we can plausibly enjoy at least most of the relevant riches of social relationship without self-deceptive belief. Relationships can be started and preserved as readily with commitment as with self-deceptive belief, and commitment allows us to treat our friends in a suitably favorable fashion without relying on belief.

Note that I am not claiming that irrational/self-deceiving belief never occurs in social relationships in the ways that James, Meiland, and Keller describe. I do think, as I will argue below, that commitment offers a better account overall of what we do and expect from our friends, but it is undeniable that many people believe and act in a self-deceptive fashion in these contexts. What I am arguing is that such practices are not demanded by the norms of friendship—that the conflict between epistemic norms and the values and norms of social relationships can be avoided. I am not attacking the psychological claim that we sometimes (or often) self-deceive about these matters, only the claim that it is necessary to do so in order to enjoy the goods of social relationships.
Is there a place for influence of friendship upon one’s beliefs? Will we always be able to sift out the influence of friendly affection from an accurate perception of the evidence? In some cases, probably not. Someone might be incapable of hearing his friend’s mistakes in her performance, or at least incapable of recognizing them as mistakes (although he might well be able to point them out while she practices.) There is something charming about the person who sincerely and vastly overestimates the talent or performance of his friend or lover or child. What I want to suggest is that we do not need to be that person in order to enjoy the goods of friendship or love or family. The love which can see the faults of the beloved is not necessarily inferior on that account. 88

8. How is this not self-deception?

Skeptics might feel that the examples described above still sound a lot like self-deceived belief. I am claiming in each case that the person does not believe, but is simply trying to have an attitude which makes them act and appear quite similar to how they would if they believed the proposition in question. What reasons do I have to offer for why we should not think there is self-deception going on?

First, as suggested in the preceding section, I would argue that an appropriate analogue/comparison is that of positive or confident thinking, which I think clearly does not have

88 "...the best love is an attentive love, a love that sees the beloved...clearly and fully, ‘as she really is,’ a love that sees the beloved’s faults and weaknesses as well as her virtues and strengths, and loves unreservedly nonetheless.” Susan Wolf, “Loving Attention,” 15.
to involve self-deception. Before an exam or performance, it helps to have a positive or confident attitude going in, but that does not necessarily rise to the level of belief. We try to dwell on considerations which suggest we will do well, try not to think about those which suggest we won’t, imagine ourselves succeeding, and so forth. But despite all this we might still be able to give an accurate appraisal of our chances and abilities if prompted to do so, although we would likely be irritated if thus interrupted in our efforts to “get in the zone” mentally.

Similarly, in the case of Michelle and Sarah, we can easily force Michelle out of her hopeful and positive mindset and into a more coldly objective one if we asked the right sorts of questions. It is hard to see how we could do this if she were self-deceived.

A looser but more colorful analogy is that of a sort of situation often found in action movies, where a character has to inch along a window ledge, or leap across a chasm, in order to escape from some villains. The important thing, of course, is not to look down. But not looking down does not require deceiving oneself about how high up one is—it is simply a matter of trying not to think about it, trying to not call it to mind. It is not that I make myself believe I am not 20 stories up, it’s rather that I try to avoid letting the 20-stories-up-ness of my situation strike me as profoundly as it does when I can see the tiny cars in the street below. This is not quite like the commitment case, as it seems a purely negative matter. But I take it that it helps illustrate that we can usefully control how we think about our situations without self-deceived belief.

A second reason for thinking that the commitment cases described above need not involve self-deceived belief is that when appropriately pushed, we demonstrate awareness of the facts of matter, even when we are unlikely to openly admit them.

If John ends up in a conversation with others who are speaking critically of Susan’s performance, and his rather feeble case in her defense is shot through of holes, and he is pushed
to say something negative, he might well throw up his hands and say “Look, she’s my friend.” He need not keep plugging away in insisting upon the excellence of her performance. He had better not beat too hasty a retreat, but his behavior is nothing like that of one who is deceived, who does not know what the facts are.

Similarly, suppose that the stakes of John’s judgment are quite high. Someone in charge of selecting the players for an elite chamber ensemble missed Susan’s performance and wants to know what John thought of it, perhaps suggesting that his decision whether to include her rests largely on John’s view. Here, I think, John hedges, and notes his friendship with Susan. If pressed further for an objective appraisal, he either answers reluctantly, or refuses to answer at all. He might do this for any number of reasons: he may not want to embarrass himself and allow others to think he has no taste or knowledge of the subject, he may want to protect the ensemble from having an unqualified player, or he may want to spare Susan humiliation. But in any case, his behavior differs from how he would act if he were self-deceived.

Similarly, if he is asked by Susan herself how he thought she did, he probably starts out very enthusiastically and approvingly. But suppose she presses him further, demanding to know what he “really” thought. Don’t we often squirm and try to avoid responding when our friends press us to tell them what we really thought about their performances or their chances of success, when they tell us to speak objectively and “not as a friend?” If friendship were truly affecting our beliefs and not our attitudes, it seems we would not squirm in that way. If we held a biased, self-deceived belief to the effect that our friends’ chances or performances were excellent, we would continue to reiterate our initial glowing perspective wholeheartedly. What we do instead is quite different—we try to assess whether our friends can handle the news or not. If they can, we might try to (very gently) speak more objectively; if we think they cannot take it, we repeat
our initial approbation, but it is now a rather conscious sort of lie. (I take it that my initial praise for my friend, even if excessive, is not really a lie, because both my friend and I know that such remarks should be taken rather lightly. Only when she pushes me to tell her “what I really thought” is there a danger of misleading her.)

A third sort of reason for thinking that commitment does not require self-deception is that our abrupt changes in behavior in certain circumstances suggest that it has more to do with non-belief attitudes than with belief. Specifically, we can switch attitudes very quickly when our friends’ actions come to affect closer friends or ourselves to a degree that we consider intolerable. I am not much bothered or affected by rumors that my friend is a thief, but if I notice some of my own possessions missing after he has visited, my attitude changes very quickly, and suddenly all the rumors of theft have a great deal of influence on my attitude. I might not take reports very seriously that claim that my friend treats women he dates badly, at least when the reports involve only strangers, but once I hear from my closer friend that he has treated her like dirt, the whole weight of the past reports comes bearing down on my attitude towards him. Even the wife in Meiland’s example, if her husband’s philandering behavior continues for quite some time, might one day decide that enough is enough, and when she does so the whole weight of the past evidence informs her outraged and disgusted attitude.

If what was going on in these cases was self-deception, something which affected our beliefs, what we took to be true, I take it that such shifts would be much more subtle. The new information would just be one data point in the balance of all this other evidence. What instead happens seems to me to be much more striking, more suggestive of an attitude shift rather than a belief shift. I try to have a particular positive attitude toward you until it affects myself or those I care about in a way I just will not stand for, and then I abandon my efforts to maintain that
attitude and have little trouble believing and acting on the negative information I was presented
with in the past.

One final reason for seeing these sorts of cases as involving an attitude that falls short of
belief is that self-deceived belief must cover its tracks in order to work, because belief involves
taking something to be true. If I am to get myself to hold an unjustified belief about my friend, I
must come to think that the belief has nothing to do with the friendship. To the extent that I am
aware of the influence of the relationship on my belief, I cease to hold the belief. Thus, to insist
that the friends in these cases are believing the relevant propositions is to insist that they are not
aware of what is going on.

But this is not actually how most of these cases work. We are often aware of how our
loyalty to our friends affects our thinking, and so our attitudes are not really full-fledged beliefs.
We often know how our thoughts, feelings, and actions would differ if we were not in this
particular relationship. If the attitude is really one of self-deceived belief, something has gone
wrong, because we are not well deceived. My friend is accused of doing something awful; I try
to take the attitude toward him that I would have if I believed he couldn’t have done whatever it
is. I need not really believe he couldn’t have done whatever it is. I might well know that human
beings are capable of all sorts of things, and even if I strongly feel like he could not have done it,
and cannot imagine him doing it, I am perfectly aware that the friends of those who do terrible
things often feel the same way. In order for my attitude to be one of self-deceived belief, it
seems that I would need to think I am special, that while other people have this tendency to have
biased perspectives on their friends all the time, I am different. Do I need to think I am special—
that I, my friends, my knowledge of my friends, and my relationships with my friends, are so
unique—simply in order to stick up for my friends? It does not seem so.
9. Why is Commitment Superior?

Commitment’s superiority to self-deception lies fundamentally in its transparency and self-awareness. To believe something I must think that I have arguments or evidence or reasoning to bolster it, however frail and fallacious. But I need not do this with a commitment. A self-deceptive belief must cover up its tracks in order to be effective. It must present itself as a well-justified belief. But a commitment need not do so.

Consider the tendency we have to think of those who are close to us in terms of unjustified superlatives. A parent may think of his child as the most beautiful in the world, yet if challenged on it, he need not get dragged into a debate about the absolute aesthetic perfection of his child’s facial features. He is unlikely to start spinning out arguments in defense of the child’s superiority over all others. He knows it was not a conclusion he arrived at through evaluation of all the evidence. He knows, in fact, that it is a commitment—that he thinks the way he does about his child because of the relationship between them, and that it is appropriate (within certain bounds) for him to do so. Or consider how we often (especially in greeting cards) tell those close to us that they are the best [mother, father, son, daughter, sister, brother, wife, husband, partner, or friend] that anyone could have. Often we are not lying or being insincere when we say such things—yet they do not rise to the level of full-fledged beliefs either, because we do not feel that they need to be supported by evidence. It would be bizarre for me, upon being presented with a story of an exceptional father, to feel the need to downplay that father’s virtues and argue for the superiority of my own father’s. Not only would it be bizarre, there would be something deeply wrong about it, as it would be a diving headfirst into self-deception. If my
attitude is one of commitment rather than self-deceptive belief, however, I can think of those close to me as “the best” without any need to create an illusion of objective evidence.

Because commitments do not hide their nature in the way that self-deceptive beliefs must, if someone challenges me about the morally significant consequences of a commitment, I can reflect on my commitment and question it. For example, if I am patriotically committed to some propositions about the goodness of my country, I can revise my attitudes if it is pointed out that my support of my country is not only unjustified but destructive. But if someone challenges me about the morally significant consequences of a belief, self-deceptive or otherwise, I will be much less responsive. If I think I hold the belief because it is true, then it will not matter much to me what the consequences of my holding the belief are.

Consider an example similar to Meiland’s: Suppose I consciously commit myself to the proposition that my friend is not stealing. But if someone warns me that he is stealing and financing some evil project with the funds and that something must be done about it, I can come to question my commitment quite readily. If, in contrast, I come to believe it self-deceptively, creating an evidence-framework to “justify” and support it, then I will have a harder time responding, because this warning will count at best as only one bit of information in the balance against all the “evidence” and self-deceptive reasoning I have been marshalling to support my belief. I do not mean that there will be no resistance in the commitment case, but it seems that it will be of an entirely different (smaller) order of magnitude. We saw this in the case of John and Susan—if it seems that overrating Susan’s performance would lead to (further) humiliation for her, John can abandon his commitment quite readily.

If the stakes of a particular commitment become too high, we can question and abandon it with relative ease. Suppose I commit myself to the proposition that I will definitely succeed in a
particular endeavor. Optimism and confidence enhance my performance, so such a commitment seems acceptable, even though the evidence only justifies a belief that I am say, 70% likely to be successful. I am aware of this to some degree, but because of the nature of commitment, I do not think about it much. But suppose I then discover that if I attempt but fail in this endeavor, someone will die as a result (we may assume that earlier I was under the impression that the stakes were trivial.) Because it is a commitment, and not a self-deceptive belief, I can then decide not to try whatever I was considering because of the intolerable risks. But if I have managed to deceive myself into believing that success is certain, I may not be able to respond appropriately to the news about the increased risk.

When a belief of ours is challenged, we respond with evidence and arguments to support it, or at least try to parry those directed against it. Not so with commitments. We can explain and justify a commitment independently of the evidence, and hold it in that way. We can admit it and honor it for what it is, without self-deceptive squirming. Self-deception, after all, is like any other kind of lie, in that it paves the way for further lies.\textsuperscript{89} Trying to self-deceptively "justify" one irrational belief spawns innumerable others. We usually cannot protect one unjustified belief and maintain epistemic integrity elsewhere.

Furthermore, the contagious nature of self-deceptive belief would often lead to unfairly forming negative beliefs about others. Stories and situations in which our friends might look bad often involve others as well, and convincing ourselves that our friend is in the right will likely bear on what we come to believe about others. For example, if I want to believe that my friend is the best on her softball team (when she is not), I must deceive myself about a great many

\textsuperscript{89} Bernard Williams, "Deciding to Believe," in \textit{Problems of the Self} (Cambridge: Cambridge University Press, 1976), 151: "Perhaps, further, one objection to the projects of believing what is false is that there is no end to the amount you have to pull down. It is like a revolutionary movement trying to extirpate the last remains of the \textit{ancien régime}.”
things—not only about her own abilities, but also about the ability of the other players, who was
to blame for various mistakes, what counts as good softball playing, and so on. Similarly, if
Meiland's Jones wants to convince himself that Smith is not stealing the funds, it seems he must
come to believe some alternative hypothesis—that someone else stole them, that the funds were
not stolen but the bookkeeper is simply dreadfully incompetent, etc. We are not simply treating
our friends better than they deserve; we are treating others worse than they deserve. It seems
unjust to think worse of others simply so we can think better of our friends.

10. Is Commitment enough for Friendship?

Even supposing that commitment as illustrated in section 7 is possible, that it does not
involve self-deception, and that it is superior to self-deception in certain respects, one might well
wonder whether it is enough to meet the demands of friendship.

Keller, anxious to show that nothing short of biased belief will do, criticizes what he calls
“pretend approval.” Rebecca does not want Eric to just say he thinks her poetry is good, she
wants him to believe it. It is true that there is such a thing as mere faking, and faking in itself is
not sign of friendship. There are all kinds of reasons why I might try to have a particular attitude
toward you. I might simply want to deceive you for my own purposes, fearing what you would
do to me if you knew the truth. If we are to imagine Eric hating her poetry reading throughout,
thinking it absolutely awful, and flatly lying simply to appease her, then Keller is undoubtedly
right that this is not what we want from our friends.

90 "Friendship and Belief," 335-6.
But I would suggest that there is often more to commitment than that. In many cases we control our attitudes holistically, out of sympathy and solidarity with our friends, and it seems to me that this can make commitment an acceptably authentic response on the part of one’s friend. Suppose that Eric throughout Rebecca’s reading truly wishes and hopes for her poetry to be good (just as Rebecca hopes and wishes for herself), and that that is what motivates his efforts to listen charitably, to dwell on the positives and pay less attention to the negatives. Suppose that attending to the negatives pains him because he cares about Rebecca and wants her to do well. It seems to me that this Eric is a good friend, despite his failure to generate the self-deceptive belief. It is not quite the same as approval, of believing the poetry to be good, but it nonetheless demonstrates a love, a loyalty, and a commitment to Rebecca as a person.

It is true that believing against the evidence does appear to be a more difficult and impressive accomplishment than this commitment, so that it seems that the self-deceiver has done more for the sake of his friend than one who merely commits. We might think that Meiland’s woman who brings herself to believe (self-deceptively) that her husband is not cheating on her is demonstrating a degree of loyalty in creating and maintaining that belief superior to that she would show by simply committing to act and think as though he is innocent.

What I hope that the earlier parts of this chapter have shown is that while such an accomplishment might be more demanding or impressive, it is not clearly better. Lying or stealing for a friend at significant risk or cost to one’s self could also demonstrate a striking degree of loyalty and love, yet it would not make those acts commendable.

In fact, I think commitments give us a better account of what we owe and expect from our friends than self-deceptive belief. What Joey really expects of Chandler in Keller’s first case is that he has a certain attitude toward Joey and Joey’s prospects. Just as Joey is presumably
dwelling on happy, hopeful thoughts about this turning out to be his big break, he expects the same of his friend, a sort of solidarity of heart and mind. If I am someone’s friend, then I ought to be committed (in the right sorts of contexts) to propositions to the effect that they will succeed, that things will turn out well for them. (This is in part because our friends are especially vulnerable to and easily influenced by the attitudes we take toward them, how we are viewing them.) I would suggest that the question of belief does not really arise for Joey or for Chandler. The optimistic mentality of those entering into a difficult contest need not be a matter of irrational belief, ignoring the evidence which points toward the possibility or likelihood of failure—it can be rather a determination to dwell on and cultivate one’s thoughts of happy possibilities, trying to maintain a particular attitude. The same can be true of their friends.

It seems irrational of the friend to demand or expect belief. If believing is as involuntary as it seems to be, how can Joey blame Chandler for not believing this job would be his big break, or how can the poet hold the friend that comes to hear her responsible for not expecting a great reading of the sort that would attract a literary agent’s attention? One wants to ask: What did these friends want? What were they expecting? If the belief that their friends have is supported by the evidence, where and how were the friends expected or obligated to deviate from the dictates of rationality for the sake of the friendship?

Consider also that in both cases that the people with allegedly epistemic expectations of their friends have their own doubts. Joey is clearly somewhat unsure of the role his new acting job will play in his future; otherwise, his sensitivity to Chandler’s response is hard to understand. Rebecca wants Eric’s supportive and sympathetic presence at her reading presumably because

---

she is feeling somewhat unsure of herself. Do they really fault their friends for not having stronger beliefs than they have themselves?

It is more plausible to interpret friendship as putting demands on the attitudes our friends should have toward us. Just as someone attempting something difficult entertains hopeful and optimistic thoughts and blocks from his mind more negative and discouraging thoughts, we expect the same of our friends. We expect them to stand in solidarity with us in our pursuit of our projects. We care because our view of ourselves and our situation is readily shaped by their view of us, by the attitude they are taking toward us. This is why we can be reassured by our friends. If it were primarily an epistemic matter, how could we be reassured? Presumably Chandler does not have better evidence than Joey does about whether this will be Joey’s big break, and presumably Eric does not have better evidence than Rebecca does about the quality of the reading that goes on at this particular venue. We are looking for someone to “be on our side,” to share our commitment with us. We are not consulting experts for information.

The final point I want to make is that the defenders of believing against the evidence in our social relationships assume that concerns with belief and what we think to be true play a far greater role than they actually do. They suggest that in every case of social interaction or doing something for a friend, I am concerned with having beliefs, a view about the truth of the matter.

Yet I would say it is unclear what my beliefs are in the case of my expectations or evaluations of a friend’s performance in a field where I have no other significant investment or interest. More importantly, the beliefs (the attitudes to certain propositions we think justified by the evidence) are simply not relevant to me. I know that I am casting my friend’s chances or performance in a favorable light, while the accurate, evidentially-justified evaluation is of no concern to me. I am simply not very concerned with the facts, with the aesthetic truth about the
performance. I neither feel the need to draw my attention to the facts to discern the truth of the matter, nor feel the need to create false beliefs about them self-deceptively, hiding the facts from myself. I do not observe my friends’ performances as a judge would—I do not watch them with a coldly objective eye, and I am not concerned with assessing and evaluating them, for the most part. It is not that I watch them in that way and then try to deceive myself about what I see.

This is even more apparent in the case of a man who wants to continue a relationship even in the face of mounting evidence that his friend is dishonest. If he is so committed to the relationship that he is willing to deceive himself in order to believe his friend to be honest, why should we think that he cares about his beliefs—what he thinks to be true—at all in this matter? Similarly, in the Jamesian cases of building up confidence to initiate a relationship, isn’t it plausible that we just aren’t that concerned in these cases about beliefs, about what we take to be rationally justified for us, about what we think to be true?

I claim that James, Meiland, and Keller all overemphasize our interest in belief and its importance to us. In a great many cases, especially those involving social relationships, we often do not care about what our beliefs are or whether we have formed them at all. Of course, we can come to care if it turns out we have good reason to—that was the point above about how commitments are more responsive to what is at stake.

But aren’t committers then guilty of an irrationality, albeit somewhat different from the self-deceivers’ irrationality? They do not sever the connection between evidence and belief, as the self-deceivers do, but don’t they instead sever the connection between belief and action, between theoretical and practical reason? Just as one law of rationality says that one’s beliefs must be based appropriately on evidence, doesn’t another say that one’s attitudes and actions should be appropriately based (in part) on one’s (evidence-grounded) beliefs.
There may be some practical cost to this, but there are practical benefits as well. If it is true, as James, Meiland, and Keller suggest, that goods such as friendship and love are lost if we insist upon applying a narrowly rational standard to our believings and doings, then it may well be broadly rational not to apply such a standard. We may have reason not to demand too much in the way of reasons in some aspects of life.

Perhaps commitments are an analogue to the partiality toward certain persons that is permitted in some non-utilitarian ethical accounts, like Bernard Williams’ talk about “one thought too many” regarding the man who concludes that his ethical duty permits saving his wife before actually doing so. We cannot imagine going through our moral lives with no consideration of ethical principles, but there are cases in which applying them is simply an instance of “one thought too many”. Similarly, one might suggest, while we cannot conduct meaningful, rational lives without epistemic principles demanding support for our beliefs, guiding what we believe and what we do as a result, in some cases, perhaps, reflecting on the evidence for a particular proposition or on even whether we believe it is one thought too many.

References


