

EPISTEMIC AUTHORITY AND ITS CRITICS

LINDA ZAGZEBSKI

University of Oklahoma

The purpose of *Epistemic Authority*¹ is to defend the rationality of belief on authority from the modern assumption that the ultimate authority over the self is the self. More specifically, the strategy is to show that traditional epistemic authority, including the teaching authority of institutions such as the Catholic Church, can be justified even if we accept the modern value of autonomy.

The book begins with a brief historical and philosophical investigation of the rejection of epistemic authority, particularly as it arises from the ideal of epistemic self-reliance. I argue that although that ideal dominates much of contemporary discourse, it cannot be defended from the work of Plato, Descartes, Locke, or Kant, and that the epistemological and moral arguments for self-reliance are weak. I then turn to an extended argument that the conscientiously self-reflective person is committed to authority in the realm of belief. I argue that epistemic (and emotional) self-trust is both rational and inescapable, that consistent epistemic self-trust commits us to trust in others, and that some of those others satisfy conditions for epistemic authority modelled on Joseph Raz's well-known theses of political authority. I apply epistemic authority to authority in communities, defend epistemic authority in the domains of morality and religion, and argue that the account of epistemic authority I give shows that epistemic authority is compatible with intellectual autonomy. Believing on authority is a demand of conscientious self-governance. It is not only compatible with autonomy, but follows from it.

¹ Linda Trinkaus Zagzebski, *Epistemic Authority: A Theory of Trust, Authority, and Autonomy in Belief* (Oxford: Oxford University Press, 2012). References to my book will be bracketed in the text.

In this essay I will give a summary of each book chapter, and will make brief comments on the critical essays most relevant to that chapter. For a few essays (Leszczynski, Anderson, and Benton) my comments are divided between two chapters. My main purpose is to give a guide to the essays in this issue, not to give a full response to each one. The authors raise a host of interesting questions on a wide range of topics pertaining to epistemic authority, religious authority, trust, and disagreement. I am grateful to the authors of these essays and to the editors of *EJPR* for this special issue. I hope that I will be able to have conversations at leisure with the authors, and that these essays will stimulate further work on epistemic authority.

CHAPTER 1. THE REJECTION OF EPISTEMIC AUTHORITY

In this chapter I trace the source of the rejection of epistemic authority to the modern rejection of authority in general, which has roots in events in the early modern era, as well as philosophical roots in the idea of autonomy. My particular interest in this chapter is the tendency to identify epistemic autonomy with epistemic self-reliance, which I find lacking in both historical accuracy and philosophical justification.

John Cottingham's beautifully written, generous and insightful essay begins with some observations on this chapter. He agrees with me that a defence of epistemic self-reliance cannot be found in Descartes or Kant, and I thank him for pointing out that we get a strong rejection of epistemic self-reliance in Wittgenstein. Cottingham then raises some important questions about moral and religious authority. The radical historical contingencies of the development of our moral systems ought to make us sceptical of the use of our own conscientious reflection in judging the trustworthiness of moral and religious authorities, and he says that he thinks my project is incomplete unless it moves beyond epistemology to metaphysics and gives an account of moral truth. Otherwise, the authorities to whom we refer can easily have sensibilities that suffer from the same historical contingencies as our own. I agree that this is a serious danger, but in response, I would say that some of the wisest persons who have ever lived, including Jesus, Socrates, and the Buddha are still identifiable, and structures have been developed to protect their insights from the contingencies of each passing age. My account assumes that rationally self-reflective persons can recognize

wisdom even when it appears in different historical periods. Traditions that developed around these figures preserve wisdom, and operate as a counterbalance to the vagaries of cultural change. One of the functions of authority in a religious community is to protect the wisdom of the community from changes arising from historical accidents. Of course, as long as we are human, we are fallible, but I do not think we have an alternative that is any better than trusting the wisest among us, and our ability to identify them. Of course, we are on firmer ground if we can rely upon divine providence, but I was not appealing to divine providence for the audience of this book.

I am grateful to Damian Leszczynski for bringing a different philosophical background to bear on the issues of this book. According to my reading of his remarks, his views are much more congenial to my theses than he implies, and in any case, he probes some important issues of philosophical methodology and the interpretation of Descartes that are helpful to an understanding of the historical background on epistemic authority. I think Leszczynski's discussion of Descartes aids my theses in two ways. First, he agrees with my claim in Chapter One that Descartes is not a defender of self-reliance, and second, he argues that Descartes is a defender of self-trust, the importance of which I stress in Chapter Two. Descartes' grounds for requiring self-trust are not the same as the ones I use in the book, but it is valuable to be reminded of Descartes' argument of the need to be certain that God is not a deceiver. Both atheists and theists need self-trust, but for different reasons. Leszczynski says, 'I could have a real objective knowledge if I could correctly apply the method, but the method is reliable only when God is not a deceiver and when we were not created by accident.' Here again I think it is useful to consider the audience. What should we say to those people who believe we appeared through a purely naturalistic mechanism? I am arguing that even if we use a method that is neutral on that issue, we get a traditional conclusion about authority by unravelling the implications of reflective self-consciousness. I am afraid, however, that there might be a misunderstanding about the method I endorse. I do not mean to be appealing to common sense. However, I assume that we have access to the structure of rational consciousness, and that it is natural in the sense that we all have it. If the structure of human consciousness inevitably leads us to the objective source of our being, so much the better for my project.

CHAPTER 2. EPISTEMIC SELF-TRUST

In this chapter I begin with the reflective self governing itself in such a way that it has two aims: to make its psychic states fit their objects, and to make its psychic states fit each other. Both of these aims are natural. The first aim includes the aim to have true beliefs and fitting emotions. The second aim is to have a harmonious self, one without dissonance. The problem of epistemic circularity shows us that ultimately, the only test that we have succeeded in the first aim is that we have succeeded in the second. Since there is no guarantee that we have succeeded in the first aim without succeeding in the second, basic trust in our epistemic faculties is inescapable. Self-trust is rational because everything we call a reason for belief is derivative from what we do when we conscientiously reflect with the aims just given. Conscientious self-reflection is the basic norm of self-governance. Since self-reflection is natural, being rational is doing a better job of what we do naturally.

Pritchard and Ryan ask whether there is empirical evidence for my claims that we have natural desires, and that among those desires are a desire for truth and a desire for a harmonious self. They are also sceptical of the way I connect the natural and the normative. They propose that the way we naturally form beliefs and respond to conflicts within the self is best explained as part of an evolutionary story, and it would be odd if that coincided with the way we ought to behave. It is not part of my project to explain how evolution connects our cognitive faculties with the way the world is, but that is clearly a problem for others whose work I welcome. I do not think that what I am doing in this project awaits the outcome of work on the evolutionary theory of the mind.

Pritchard and Ryan then point out that I have not solved the problem of radical scepticism, but have shown that according to the view of rationality I propose, it is less rational to be a radical sceptic than to be self-trusting. I agree with them that that is what I do. I have no intention of resolving the sceptical problem, and I agree that what makes scepticism so worrisome is that it seems to be the consequence of what we do naturally—thoroughly seeking reasons for our beliefs. As I say in discussing Alston (p. 41), the person who desires full reflective justification for her beliefs and tries to attain it is doing what every reflective person does, only more thoroughly and scrupulously. When she does so, she finds that she is attempting the impossible. The issue for me is what response is most rational, and I give my reasons for thinking

the self-trusting stance is the most rational, but I have no solution to the paradox of scepticism, nor to Alston's paradox of the desire for full reflective justification.

Matthew Benton raises a related question in the section of his essay on rationality and the resolution of dissonance. I describe the situation of the reflective person who becomes aware of the problem of epistemic circularity. She trusts her faculties in her pre-reflective state, but when she realizes she cannot attain what Alston calls 'full reflective justification' for her beliefs, she must either achieve a higher level of trust, or she can live with the dissonance. I say that roughly, rationality is doing a better job of what we do anyway. Benton says he can see why trust after encountering the circularity problem is doing what we do anyway, but why isn't it doing a better job if we lose trust, given that we discover upon reflection that we lack something we rationally want? Alternatively, we could live with the dissonance since, as I point out, we can live with some dissonance. Not all dissonance has to be resolved.

I thank Benton for pushing me in the direction of living with dissonance. This is a point I wish I had made in the book. I do not want to say that a rational person must give up the desire for full reflective justification. She may not be able to do so. Often desires continue long past the time we realize they will never be fulfilled. But the issue for my project is what does she do about her pre-reflective trust? She must live in a psychic world in which it is impossible to satisfy a certain desire. If her psyche depends upon the satisfaction of that desire, then she will do what the radical sceptic (allegedly) does, and the structure collapses. But reflective trust permits her to continue managing her psyche as she always did, but in a more realistic way because now she knows that she has a desire for the impossible. But what if she does *nothing* in response to the awareness of circularity? Suppose that nothing changes in her pre-reflective trust in herself, and nothing changes in her beliefs or desires. She does nothing about it at all. She just doesn't think about it. I think that is the person Benton has in mind in one of his remarks. Can she live in a permanent state of dissonance without changing anything that generates the dissonance? It would be interesting to know whether such a thing can happen. I suspect something changes unconsciously. Either she gradually and with no conscious awareness becomes a sceptic, or she gradually and unconsciously realizes the critical function of self-trust in her psyche and it rises to the reflective level. But perhaps not. I don't think I can say what it is rational to do without knowing more about

what people actually do. This is an interesting possibility and I thank Benton for mentioning it.

Leszczynski also has some interesting comments on dissonance, and I like his use of Socrates to illustrate the advantages of experiencing dissonance in our psychic states. Because our consciousness is not insulated from the world around us, including the social world, we cannot get away with living the life of a coherent paranoid. We must continually adjust our conscious states – beliefs, emotions, etc. – in response to experiences of that world. I agree that dissonance is a good thing because it forces us to respond to it by changing something in our psychology– a belief, a putative memory, an interpretation of an experience, an emotion, and so on, and one of the reasons we need self-trust is that we think that in doing so, we are getting closer to having a mind whose states fit the world they are about– true beliefs, fitting emotions, veridical memories. I think that that is a good way to understand what it means to be rational.

CHAPTER 3. EPISTEMIC TRUST IN OTHERS

In this chapter I argue that consistent epistemic self-trust commits us to the same kind of trust in others. Two forms of epistemic egoism are incoherent, and we are committed to a weak form of epistemic universalism. The fact that another person has a certain belief always gives me a *prima facie* reason to believe it. If another person holds a belief conscientiously, I have a stronger *prima facie* reason to believe it. I distinguish two kinds of epistemic reasons: first person or deliberative reasons, and third person or theoretical reasons. Trust in self and trust in others are deliberative reasons for belief. These reasons do not aggregate with third person reasons, or what is often called evidence.

Charity Anderson's essay is delightfully clear and rewarding to read. In her discussion of this chapter she correctly observes that my universalist principle is weak. The fact that someone else has a belief *p* always gives me a *prima facie* reason to believe *p*, but the reason is not decisive and it can be defeated, perhaps even easily defeated. Nonetheless, it always has some weight. Anderson worries that the cases in which this principle has an effect on our beliefs are few, and that that reduces the effectiveness of the argument to oppose the egoist. She makes the parallel point about my argument that we owe a higher degree of trust

in the faculties of conscientious other persons whose conscientiousness we discover by being conscientious ourselves. Most people are not conscientious most of the time, and so, she argues, the principle does not commit us to very much.

I am neutral on the number of cases in which a universalist would believe what some other person believes and an egoist would not. But even when the universalist and the egoist coincide in believing or not believing some other person, I think that there is a world of difference between the person who approaches the epistemic faculties and beliefs of others with the attitude 'Innocent until proven guilty', and someone who approaches other persons with the attitude 'Demonstrate your innocence', or even worse, 'Guilty until proven innocent'. The first is like the attitude of the epistemic universalist, the second is like that of the standard epistemic egoist, and the third is like that of the extreme epistemic egoist. I think that universalism not only can make an important difference in certain critical cases of belief, but it forces us to become intellectually humble. A disposition to see others as basically equal to ourselves, and hence partners in the search for truth, can mitigate our natural tendency to inflate the self, and to close our mind and heart to the riches of other minds and the results of their reflective efforts to get the truth. Anderson is right to mention that self-deception can lead us to misjudge our own degree of conscientiousness, and I think that the trust in others to which our self-reflection leads can also have the effect of making us more realistic about our own defects in the use of our powers.

In Chapter Two I argue that ultimately, our only test that a belief is true is that it survives future conscientious reflection. That is because the problem of epistemic circularity means that we never have a final verdict that a belief is true, and there is always a possible gap between our evidence or reasons for belief at any given time and the truth. But we think that if a belief is true, it will not be disconfirmed by future evidence. Since I argue that many different kinds of psychic states can function as deliberative reasons for a belief, ultimately, the only way I have to tell that a belief is true is its survival of my conscientious reflection now and into the future when my experiences change, and I have given up some of my other beliefs. So the future confirms or disconfirms what I believe now. To believe *p* is to think *p* is true, and to think *p* is true is to make a bet on our future reasons for belief. I think we make these bets all the time. If what I believe now is disconfirmed by future evidence, then I will have to change my mind, and there is nothing wrong with

that. Change is a necessary part of the process of making our beliefs closer and closer to the truth. But sometimes I can be confident that the future will not disconfirm my present beliefs. This is particularly true for beliefs that are central to the self, such as moral and religious beliefs, and beliefs closely tied to one's personal commitments. I think that this is important for the resolution of disagreement with other persons. I do not offer a formula for resolving disagreement, but stress the fact that it is a dilemma that arises within the consciousness of each individual conscientious person. Anderson says that we will not always know the conscientious way to respond, but that is as it should be if it is resolved by a self-directing person. Granted, it is hard to predict what will survive our future reflection because we do not know what future experience will bring, but to some extent we decide our future self. Each person has to decide where she will hold the line – 'This belief is part of me' – and where she will leave it open that the belief may have to be given up.

CHAPTER 4. TRUST IN EMOTIONS

This chapter gives arguments for trust in emotions parallel to the arguments of the previous two chapters. We have the same kind of reason for basic trust in our emotion dispositions as we have for basic trust in our belief-forming faculties, and we have the same kind of commitment to trust in the emotions of others. This makes emotion disagreement a problem parallel to the well-known problem of epistemic disagreement. Trust in admiration gives us another route to trust in the *beliefs* of others. Intellectual admiration can be a deliberative reason to trust the deliverances of the epistemic faculties of admirable persons.

CHAPTER 5. TRUST AND EPISTEMIC AUTHORITY

This chapter moves from trust in the beliefs of others based on their similarity to ourselves to trust in the beliefs of epistemic superiors, persons that we conscientiously judge are more likely to get the truth than we are. This leads to a defence of epistemic authority modelled on Joseph Raz's theses of practical authority. The two main theses are the Pre-emption thesis and the Normal Justification thesis. The *Pre-emption thesis* is a thesis about what it means to take someone as an epistemic authority. According to that thesis, the fact that the authority has a belief p

is a reason for me to believe *p* that replaces my other reasons relevant to believing *p* and is not simply added to them. Raz's *Normal Justification thesis* says, roughly, that the normal way to show that *A* is an authority for *S* is to show that *S* is more likely to act on his first order reasons if he does what *A* says to do rather than to try to act on those reasons directly. I have two *Justification of epistemic authority theses* that are analogues of Raz's NJ thesis. The first says that the authority of another person's belief for me is justified by my conscientious judgment that I am more likely to form a true belief and avoid a false belief if I believe what the authority believes than if I try to figure out what to believe myself. The second says that the authority of another person's belief for me is justified by my conscientious judgment that I am more likely to form a belief that survives my conscientious reflection if I believe what the authority believes than if I try to figure out what to believe myself. I argue that even the epistemic egoist should accept the first thesis. When someone else satisfies one of these theses, I should take her belief pre-emptively.

I was delighted to read **Arnon Keren's** paper on epistemic authority and pre-emption since he has already engaged with these issues. Keren agrees with my central claims, but argues that there are important differences between practical and epistemic authority that damage the line I take in defending the justification of epistemic authority. Keren agrees that Raz's pre-emption thesis applies to epistemic authority, but he argues that some of Raz's other theses of authority do not. Keren appeals to our practices of criticism to support the claim he and I share that epistemic authority is the power to generate pre-emptive reasons for belief, but he strengthens the position by claiming that epistemic authority is the normative power to make it a *duty* for others to believe pre-emptively. I prefer to refer to what a rational, self-governing person 'should' do, and do not use the language of duty, but my reasons for not speaking of duty have to do with my views about duty in general that I do not bring up in the book, and I see no reason to object to Keren's position. If there is any distance between us on this point, I do not think it is serious.

Keren does differ from me on Raz's No Difference thesis. According to that thesis, there is nothing that subjects ought to do as the result of the exercise of authority which they did not have reason to do anyway. The authority's directive merely gives them new reasons for doing what they already had reason to do. Raz argues that the No Difference thesis is false for authority in general. Keren argues that the thesis is false for

practical authority but true for epistemic authority. I argue that the No Difference thesis is false for epistemic authority as well as for practical authority. When a practical authority tells me to do X, I have a reason to do X that is not simply another reason to do what I already had reason to do. Similarly, when an epistemic authority tells me that p, that gives me a reason to believe p that is not simply another reason to believe what I already had reason to believe. I do not understand Keren's claim that the No Difference thesis applies to epistemic authority because he seems to be highlighting a different, but very interesting difference between practical and epistemic authority. In the case of epistemic authority, Keren points out, the authority cannot give me a reason to believe p unless she has one herself. In contrast, a practical authority can give me a reason to do X when sufficient reasons to do X did not previously exist. Keren is right that this is an interesting difference that deserves investigation because it applies to the current controversy about the way knowledge is transferred via testimony. It is a difference in the conditions for the exercise of responsible authority, a difference that applies to the person in authority, not to the subject. However, I do not see that it pertains to the No Difference thesis.

Keren agrees with me that epistemic authority gives me pre-emptive reasons to believe what the authority tells me, but Keren disagrees with my defence of pre-emption by referring to Raz's financial shares argument in which Raz argues that unless one pre-empts in taking the authority's advice, one's track record will be worse overall. Keren points out that in the practical case there are only two options: sell the stock or not sell the stock, whereas in the epistemic case, there are three options: believe p, believe not p, or suspend belief. As long as our epistemic goals include both getting truth and avoiding falsehood, he argues, then withholding judgment, while not the best outcome, allows us to guarantee an epistemic outcome that is second-best. So instead of pre-empting (always believing the expert), we are better off in withholding judgment in at least some of the cases in which the expert and I disagree. So Keren's idea is that I have to weigh three possible outcomes: (i) believing a truth, which is the best outcome, (ii) believing a falsehood, which is the worst outcome, and (iii) withholding belief, which means I might lose the chance for the best outcome, but at least I avoid the worst.

I think this is an excellent case for discussion of pre-emption because there are a number of different scenarios that deserve discussion. I think Keren is right that the subject has three options, and I thank him for

pointing that out, but notice that the authority has three options also. So the authority can say 'Believe p ', 'Believe not p ', or 'It's inconclusive whether p , so withhold judgment'. It will still turn out that my track record is better if I follow the authority pre-emptively under these conditions. That is, I should believe when she says to believe, and withhold when she says to withhold. But there are other scenarios that raise William James' famous point in 'The Will to Believe' about weighing the value of truth against the disvalue of falsehood. For instance, if my hatred of falsehood exceeds my love of truth, it would be reasonable for me to withhold belief more often than I would otherwise. We must, then, be clear about what our first order epistemic ends are.

Benjamin McMyler also pursues the topic of pre-emption in his essay, arguing persuasively that authority includes more than the power to give pre-emptive reasons, but also includes an essential interpersonal dimension. McMyler begins with a worry that the authority of someone else's belief is not strong enough to be parallel to Raz's view of practical authority. The reason is that Raz says that authority is exercised in the giving of commands or directives, which means that the parallel exercise of epistemic authority is the authority of testimony, which I discuss in Chapter Six, rather than the authority of belief, which I discuss in Chapter Five. I think that our disagreement here is minimal. I am willing to say that the authority of belief is weaker than the authority of testimony, which has the addition of important interpersonal features, but there is something interesting about Raz's Normal Justification thesis even if authority is exercised in the intentional issuing of directives. If someone knows better than I how to act on my first order reasons, she knew it before she told me anything. So her satisfaction of the Normal Justification thesis does not require that she testify that p to me, and that is why someone's belief can be authoritative for me, and it is also the reason that inanimate objects like GPS systems can be authoritative for me. Nonetheless, I agree with McMyler that the interpersonal dimension adds something important to authority, something that explains why it is stronger than the authority of navigation systems, and why it has often raised worries about the compatibility of authority and autonomy. When I obey someone who is authoritative for me, I am submitting to that person. Trust between authority and subject is essential to the proper operation of authority. Paradoxically, authority is an aspect of conscientious self-governance because I need others who intentionally and effectively help me govern myself according to standards I accept.

The conclusion is that McMyler is right that the normative power to give others pre-emptive reasons is not sufficient for authority in its most interesting sense.

Trent Dougherty defends evidentialism and the Lockean view that faith is belief on the evidence of the testimony of an expert. He says that the view that faith is a kind of belief on a kind of evidence is 'run of the mill'. He sees no reason why a subject of authority would take the authority's testimony pre-emptively when she already has reasons that bear on the issue. He does not comment on the Razian reason that we are more successful at reaching our own ends if we pre-empt. However, he says he doubts that it is even possible to pre-empt, and I agree that the psychology of pre-emption deserves close investigation. If pre-emption is essential to obeying authority, and if it is impossible to pre-empt, it would follow that it is impossible to obey authority. If pre-emption is not essential to obeying authority, and if the subject merely considers the authority's directive as another piece of evidence, it is hard to see how obedience is anything other than doing what a rational person would do in any case. Dougherty says he freely made a vow to obey the Church when he was confirmed, but I would like to discuss with him his reasons for making the vow since it seems to me likely that he judged that the Church satisfies some thesis of the justification of authority similar to the ones I propose. (In fact, his reasons were probably more reflective than mine were since he was confirmed as an adult, whereas I was only twelve).

The main dispute between Dougherty and me, however, is on the nature of evidence, and what it means to believe on evidence. He says that evidence is anything, broadly construed, that aims at truth. As I mentioned above, I argue in Chapter Two that since we have no guarantee that anything we call evidence (or reasons for belief) leads us to truth, evidence is less basic than self-trust in our epistemic faculties. Nothing we call a reason or evidence would be a reason unless it is reasonable to trust the connection between the conscientious use of our faculties and getting the truth. Self-trust *is* a reason, and because it is a reason, all the other reasons we can identify have the status of being reasons, and that includes everything Dougherty calls evidence. The difference is that self-trust reveals the first person states that we take to indicate truth in addition to propositional evidence. Emotions can be in that category, and that is why I say that trust in the emotion of admiration can give us another route to the justification of authority. But even if all we are

doing is weighing evidence in the Lockean sense, we get an argument for the reasonableness of taking certain persons as authoritative on the Justification theses. What I think Dougherty adds is the need to distinguish degrees of authority, an important thing to do when there are competing authorities, or when I am not sure whether someone else is so much more likely than I am to get the truth in some domain, that I should consistently, and for the foreseeable future, take their testimony in that domain pre-emptively. But I think that even though the idea of degrees of authority is an important issue, it overlooks something important about authority – the interpersonal dimension stressed by McMyler. For some authorities, I invest myself in them. I do not say, for instance, I am 70% sure of you, and only 68% sure of some other religion. As Coady argues below, it is hard to say that the authority of the Church is all of nothing, but it can't be a matter of dividing a certain percent of my loyalties between one putative authority and another either.

In the second half of Anderson's paper, she brings up some interesting questions about pre-emption. One question that I find particularly fascinating is the question of counterfactual stability. Have I acted on authority if I let the authority's directive pre-empt my other reasons, but I might very easily not have done so? I had not thought of this question before and have no immediate answer. If counterfactual stability is not necessary, that seems to mean that we can act on authority accidentally. But if counterfactual stability is necessary, the conditions for acting on authority become very strong, perhaps too strong. I hope that other writers on pre-emption have something to say about this question.

CHAPTER 6. THE AUTHORITY OF TESTIMONY

This chapter gives a parallel justification for the authority of testimony, which differs in some interesting ways from the authority of someone else's belief. I defend a view of testimony that is strongly anti-reductionist, similar to the views of Moran and McMyler.

In the second section of Matthew Benton's paper, he brings up a kind of 'authority' that is possessed by a testifier who knows what she tells me. I do not address knowledge or the transfer of knowledge in my book, and I do not treat the testifier who knows *p* as possessing authority in the sense I mean. Benton is right that the conditions for successful testimony he discusses are third personal, and the pre-emption thesis does not apply. But in this chapter I am interested in a subset of testimonial cases

that satisfy the justification thesis, and which pose some issues that are an extension of the issues discussed for the authority of someone else's belief in the previous chapter. The problem for most of the book is from the first-person perspective, although I argue for a third-person version briefly. The problem of this chapter therefore has very little in common with most of the literature on testimony.

Benton's discussion of testimony brings up an interesting issue in the last section of his paper, in which he discusses testimony as evidence vs. the trust model of testimony as applied to religious belief. Readers will find a very nice discussion of how we can know that a revelation is from God in the first place, prior to the operation of the trust model.

CHAPTER 7. EPISTEMIC AUTHORITY IN COMMUNITIES

Authority is most interesting when it involves networks of people in communities with a structure of authority that serves the purposes of the community. This chapter begins with differences between political authority and authority in small communities with a high degree of trust. These differences permit the justification of a stronger kind of authority than we get from Raz's Justification of Authority principle, and can apply to epistemic authority.

The epistemic authority of my community is justified for me by my conscientious judgment that if I believe what We believe, the result will survive my conscientious self-reflection better than if I try to figure out what to believe in a way that is independent of Us.

Tony Coady's paper is relevant to this chapter, as well as to Chapter 9, and I discuss it below.

CHAPTER 8. MORAL AUTHORITY

In this chapter I argue that the justification of epistemic authority applies to moral beliefs, whether the authority is an individual or a community. Epistemic egoism is very common as applied to moral beliefs, and even extreme egoism has a substantial number of adherents. In fact, it is probably the dominant view. I argue in this chapter that moral epistemic egoism is incoherent. Taking a moral belief on authority is justified in the same way taking non-moral beliefs on authority is justified. Moral understanding, in contrast, cannot be attained through the testimony of authorities, but even understanding can be aided by other persons.

Cottingham's paper, discussed under Chapter One above is relevant to this chapter, and Coady's paper, discussed under Chapter Nine, is also relevant to the topic of moral authority.

CHAPTER 9. RELIGIOUS AUTHORITY

This chapter argues against religious epistemic egoism, and uses the defence of religious epistemic universalism as a way to generate a novel form of the *consensus gentium* argument for theism. I discuss divine testimony and give three models of revelation: the model of chains of testimony about an original experience of the divine, the model of recurring first-person experience of the divine, and the 'high point' model in which a high point of history after an original revelation creates a tradition. I conclude with a justification of the religious authority thesis which says that the authority of my religious community is justified by my conscientious judgment that if I engage in the community, following its practical directives and believing its teachings, the result will survive my conscientious reflection better than if I try to figure out what to do and believe in a way that is independent of Us.

Three essays address issues raised in Chap Nine. John Schwenkler and Jacek Wojtysiak discuss models of revelation. C. A. J. Coady discusses authority in religious communities. Coady's comments are also relevant to Chapters Seven and Eight.

John Schwenkler discusses two problems I raise for the first model of revelation. On the chain model, nobody is as justified in a belief acquired through a chain mechanism than the first person in the chain. The farther away we are from the source, the less justified we are. Schwenkler quite rightly points out that actual religious traditions always add many features to this model that increase the justifiability of the belief: there are multiple chains, there is back-tracking and double-checking, convergence on a single person from multiple directions, and most importantly, supervision of the entire process. These features bring the model much closer to the third model I describe, in which there is a development of the tradition over time under the supervision of a teaching authority. But the feature Schwenkler adds that is closest to the high point model is what he calls Summation: Each original witness is given only a proper part of the message, which is put together when the chains converge. The third model I describe is similar to Schwenkler's up to the point of convergence. On my third model, what is transmitted is not an original

experience, but an interpreted body of oral and written traditions put together from various sources at a high point that is thereafter transmitted to succeeding generations. The transmission is intended to have continuing relevance to future members of the community, and with an authoritative method to preserve the high point of the tradition. James Kugel argues that Jewish tradition should be understood on this model, and I argued that Catholic tradition is also. As Schwenkler points out, the chain model does not explain how the transmission of tradition is improved, yet we ought to accept its improvement from other sources of truth in addition to sacred texts and ancient theological documents.

In another essay on models of revelation, Jacek Wojtysiak plausibly argues that the experience model is a refinement and expansion of the chain model, and the high point model is also dependent upon the chain model since the community's beliefs need to include a reliable story on the initiating event of divine intervention. Furthermore, the concept of a high point implies a structure in which the chain is modified to a chain of transmission of a way of understanding God. This must involve both recalling remote events, and applying the community's ancient exemplars of faith and practice to every believer's life.

Tony Coady has written a very sensitive discussion of communal authority, particularly in the way it can become corrupted when institutionalized. If I understand him, he does not object to my theses of epistemic authority as applied to communities, but argues that I apply them in a way that seems to lead to an 'all or nothing' acceptance of an authority, which is unrealistic for an institution like the Catholic Church. My model starts with the authority of an individual, then moves to authority in small communities, and I then apply the model to the authority of the Church and other large religious communities that have become institutionalized. Groups of people can recognize the same authorities, and practical and epistemic authority can often become intertwined. Somebody can be better than I am at both determining how to reach my practical goals and how to get true beliefs, and there can be general agreement about this among a group of people. Rich and varied communities can arise in this way. But when the authority gives a large number of practical directives as well as testimony to a large number of beliefs, complications are bound to arise, as Coady rightly points out. The community might split apart for what seems to us now to be a trivial matter (such as the East/West dispute over the *filioque* clause in the Nicene Creed), yet it might not split apart over more serious matters

of sexual morality, contraception, and abortion. How can that happen? My view is that there is both personal and communal self-reflection, and they do not necessarily coincide in their results. I may conscientiously come to the conclusion that some particular teaching of my Church does not satisfy my conscientious reflection upon my total set of beliefs, values, and experiences, and so the Church is not more likely to reach the truth about that particular matter than I am myself. But of course, if I make such a judgment very often, that will weaken my belief that in general the Church is more like to reach the truth (or get a belief that will survive my conscientious reflection) than I am myself. So Church members may disagree about certain doctrines and still remain members, but one of the things they may disagree about is what they need to agree about in order to remain members. This is a problem in ecclesiology. Every community has to decide what makes the community what it is, and what authority structure preserves it as what it is. The Church has a communal set of beliefs, values, historical memories, and goals upon which the community continually reflects, and its authority structure is one of the things upon which it reflects, sometimes leading to important changes, as happened during Vatican II.

Coady rightly notes that I do not discuss institutional authority, and he brings up some important problems that arise once a political dimension is introduced. He suggests that that can lead to a separation between the wisdom dimension of authority and the institutional dimension, where the persons with power in the institutional structure are not the wisest (or holiest) persons. This is an important point, and I thank Coady for pressing me on it. I think that institutional structure is itself justified by communal reflection, but when that means that those in power lead the reflection, there is a strong bias in favour of maintaining their power. However, there are always voices of wisdom, not the least of which is the Pope (and it is interesting that Pope Francis has arranged his life in such a way that he cannot be 'protected' by those around him from finding out what he needs to know, a particular problem for the previous two popes). I also think it is important that a community, whether institutionalized or not, must not succumb to what I call communal epistemic egoism (Chapter 10, sec. 4), the analogue of personal epistemic egoism, and that gives a community the obligation to conscientiously listen to and respond to criticism from the outside as well as dissent from within. I think it is also worth noting that the ideal authority structure is not necessarily democratic since democratic institutions are subject to the

same kinds of corruption lamented by Coady, and in any case, epistemic authority by its nature cannot be democratic. I realize that many people are mistrustful of institutions in general, but I think that institutions are necessary for the perpetuation of moral and religious insights. Coady is obviously right that power corrupts, but power serves some important human purposes as well.

CHAPTER 10. TRUST AND DISAGREEMENT

This chapter addresses the difficult problem of how to handle conscientious disagreement with other persons whom I conscientiously trust. Given that trust in others is a commitment of the conscientious self, and given that I conscientiously judge that some others are equal or superior to myself in their ways of getting the truth, there can be a conflict between a belief I conscientiously hold and a conscientious belief held by someone with an opposing belief who is at least as trustworthy as myself. Going back to the distinction between theoretical and deliberative reasons, I argue that the problem is not mysterious if it reduces to a conflict of theoretical reasons since there are many cases in which evidence points in conflicting directions. The real problem is a conflict among our deliberative reasons, which are irreducibly first personal. The problem can only be resolved by determining which is most conscientiously trusted. The last section of the chapter turns to communal epistemic egoism and the need to resolve disagreement between communities in a way parallel to the resolution of conflict within the self.

At the beginning of the chapter I reject two extreme positions on this issue – the egoist position that the beliefs of others do not count, and the egalitarian position that my reasons for belief count no more than the reasons of others. But I also reject compromise positions which on many issues can lack the theoretical support of either of the extreme positions. I call the disagreement problem an antinomy because I think it shares some features of Kant's antinomies, where the solution is not for each side to give a little to the other side, but to change the terms of the debate. The outcome may be that one side 'wins' in a sense, but not on the terms of the original dispute. **Joshue Orozco and Nathan King** give a very helpful summary of the 'total evidence' position of Thomas Kelly and others, which is a well-motivated middle position that is not a mere compromise between extremes, and as they observe, is not far from my own view. But since the problem in its strongest form arises from my

first-person viewpoint, Orozco and King think that I am closer to the egoist position because ‘evidence drops out’ and is replaced by survival of conscientious reflection.

At this point in their essay I think that Orozco and King have a misunderstanding of my view of rationality as survival of future conscientious reflection. The norm of reflection is not to believe what my future self will believe, but to believe in a way that survives my conscientious reflection on into the indefinite future. As I said in discussing Chapter Two and Anderson’s essay above (Chapter Three), anything we call reasons for belief derive from what we do when we conscientiously reflect with the aim of truth, but since there is never a final verdict on whether we have been successful in reaching the truth, we aim to believe in a way that survives reflection on into the future when we may have more evidence. Orozco and King describe a physicist attempting to confirm or disconfirm the existence of the Higgs Boson. She conscientiously reflects upon the data she has and forms a judgment that it exists. Orozco and King say she is rational in her belief and I agree, with one provision. If she has reason to believe that the data will differ in the future, she should form her judgment provisionally. But of course, she may not know whether or not the data will change, in which case her belief that the Higgs Boson exists includes a bet that her belief will not be disconfirmed by future evidence. As I say above, I think we make these bets all the time. There is nothing mysterious about them. But one thing she does not do is to think about what her future self will think. I agree with Orozco and King that she is not doing that. Furthermore, I say in the book that it is not even necessary that she ask herself what will survive her future conscientious reflection. It is necessary that she forms her beliefs in a way that will survive reflection, but it is not necessary to ask herself, ‘Will this belief survive conscientious reflection?’ She need not raise the second-order question about her own level of rationality very often.

CHAPTER 11. AUTONOMY

In the final chapter of the book I argue that conscientious self-reflection is the fundamental norm of autonomy. Attacks on self-trust are attacks on autonomy. The conscientiously self-reflective person described in this book is the autonomous agent. Since conscientious self-reflection shows us the rational justification of belief on authority, authority is compatible with autonomy and, in fact, authority is entailed by autonomy.