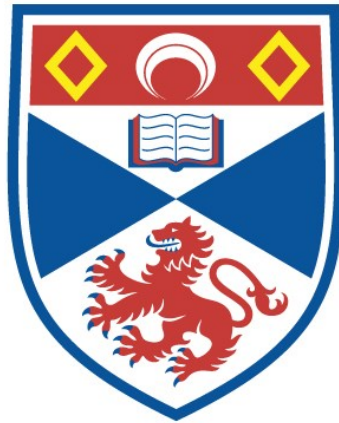


Moral inquiry: we're all in this together

Nick Küspert

A thesis submitted for the degree of PhD
at the
University of St Andrews



2024

Full metadata for this item is available in
St Andrews Research Repository

at:

<https://research-repository.st-andrews.ac.uk/>

Identifier to use to cite or link to this thesis:

DOI: <https://doi.org/10.17630/sta/811>

This item is protected by original copyright

They talk about the people, but they do not trust them; and trusting the people is the indispensable precondition for revolutionary change. A real humanist can be identified more by his trust in the people, which engages him in their struggle, than by a thousand actions in their favor without that trust.

Paolo Freire, *Pedagogy of the Oppressed*

Contents

Declarations	i
Acknowledgments	iv
Abstract	vi
Introduction	viii
1 Conciliating to <i>Avoid</i> Moral Scepticism	1
1.1 Introduction	2
1.2 Symmetry	3
1.2.1 Belief	5
1.2.2 Reasons for Belief	7
1.2.3 Permissivism	9
1.2.4 Extreme Disagreement	10
1.3 Moral Agreement	12
1.3.1 Moral Conciliationism does not entail Scepticism	13
1.3.2 Moral Inquiry as a Collective Enterprise	15
1.4 Conclusion	18
2 The Limits of Pessimism About Moral Testimony	21
2.1 Introduction	22
2.2 Moral Understanding	24
2.2.1 Moral Worth	25
2.2.2 Moral Worth Fetishism	30
2.2.3 Moderate Optimism	33
2.3 A Problem in the Agent?	34
2.3.1 Non-Integrated Moral Beliefs	34
2.3.2 ‘Cold’ Moral Beliefs	36
2.4 Intertwining Dogmas	40
2.5 Conclusion	42

3	I think I got it right	45
3.1	Introduction	46
3.2	The Parity Claim	48
3.2.1	Moral Understanding	49
3.2.2	Conative Attitudes	51
3.2.3	Autonomy	52
3.2.4	Authenticity	53
3.3	Trusting Yourself and Trusting Others	57
3.3.1	No Ulterior Motives	58
3.3.2	A More Nuanced Optimism	61
3.4	Conclusion	64
4	Moral Experts: Authorities <i>and</i> Advisors	66
4.1	Introduction	67
4.2	Moral Experts	68
4.3	Authorities and Advisors	69
4.4	The Hybrid Proposal	72
4.4.1	The Aim of Moral Inquiry	74
4.4.2	Refining the Hybrid Proposal	77
4.5	Conclusion	81
5	Moral Expertise and Moral Philosophy	84
5.1	Introduction	85
5.2	Why an Independent Check?	86
5.3	Theoretical Knowledge	88
5.4	Derivative and Fundamental Moral Reasons	90
5.5	Derivative Reasons and Moral Expertise	92
5.6	Traditional Conceptions of Moral Expertise	93
5.7	Conclusion	95
6	Moral Testimony and Blame	98
6.1	Introduction	99
6.2	Blaming You	100
6.3	Blaming Your Testifier	103
6.3.1	Passing the Buck	105
6.3.2	Optimism and Pessimism	108
6.3.3	Moral Vulnerability	109
6.4	What Kind of Inadequate Care?	110
6.5	Conclusion	111
	Concluding Remarks	114
	Bibliography	118

Declarations

Candidate's declaration

I, Nick Küspert, do hereby certify that this thesis, submitted for the degree of PhD, which is approximately 60,000 words in length, has been written by me, and that it is the record of work carried out by me, or principally by myself in collaboration with others as acknowledged, and that it has not been submitted in any previous application for any degree. I confirm that any appendices included in my thesis contain only material permitted by the 'Assessment of Postgraduate Research Students' policy.

I was admitted as a research student at the University of St Andrews in October 2020. I received funding from an organisation or institution and have acknowledged the funder(s) in the full text of my thesis.

St Andrews, 18/12/2023

Supervisor's declaration

I hereby certify that the candidate has fulfilled the conditions of the Resolution and Regulations appropriate for the degree of PhD in the University of St Andrews and that the candidate is qualified to submit this thesis in application for that degree. I confirm that any appendices included in the thesis contain only material permitted by the 'Assessment of Postgraduate Research Students' policy.

St Andrews, 18/12/2023

Permission for publication

In submitting this thesis to the University of St Andrews we understand that we are giving permission for it to be made available for use in accordance with the regulations of the University Library for the time being in force, subject to any copyright vested in the work not being affected thereby. We also understand, unless exempt by an award of an embargo as requested below, that the title and the abstract will be published, and that a copy of the work may be made and supplied to any bona fide library or research worker, that this thesis will be electronically accessible for personal or research use and that the library has the right to migrate this thesis into new electronic forms as required to ensure continued access to the thesis.

I, Nick Küspert, confirm that my thesis does not contain any third-party material that requires copyright clearance. The following is an agreed request by candidate and supervisor regarding the publication of this thesis:

- No embargo on print copy.
- No embargo on electronic copy.

St Andrews, 18/12/2023

St Andrews, 18/12/2023

Underpinning Research Data or Digital Outputs

I, Nick Küspert, hereby certify that no requirements to deposit original research data or digital outputs apply to this thesis and that, where appropriate, secondary data used have been referenced in the full text of my thesis.

St Andrews, 18/12/2023

Acknowledgments

At the heart of this project is the insight that moral inquiry is fundamentally a social enterprise. The same, I think, goes for inquiry more generally, and this dissertation is certainly no exception.

Writing this dissertation would not have been possible were it not for my patient supervisors at the *University of St Andrews*, Justin Snedegar and Jessica Brown, as well as my quasi-supervisor at *Harvard University*, Zoë Johnson King. I am grateful to all three for giving up much of their time to discuss my ideas and my writing, allowing me to turn these ideas into this dissertation. I am likewise indebted to my annual reviewers these last three summers: Adam Etnison and Kent Hurtig, as well as the late Sarah Broadie. Adam Etnison also served on the examining committee, as did Guy Fletcher—I would like to thank both of them for spending much of their winter months reading this dissertation.

It is not just the close supervision that makes academic writing possible, it's also the multitude of staff running our administration, our libraries and keeping our buildings in order. I am particularly grateful to Amy Baird, whom I miss dearly.

During my PhD, there have been people in my life whom I could share my academic and non-academic worries and joys with, and I don't think I would have been able to complete much of this dissertation without them. Jakob Eichler, Camille Fouché, Enrico Galvagni, David MacDonald, Thomas Randriamahazaka, Iza Regliński, Luca Stroppa and Zara Treharne: thank you so, so much, for your friendship.

Last but not least, I am indebted to my family. Were it not for them, I couldn't have finished (let alone started) this dissertation. Mama, Papa, Tom and Julie—thank you for your love.

Throughout my enrolment in the PhD programme at the *University of St Andrews* and the *University of Stirling*, I received a generous stipend from the *Philosophical Quarterly* as well as travel bursaries from the *Department of Philosophy* and the *Arché Philosophical Research Centre* here in St Andrews. My stay at *Harvard University* was possible only due to its fellowship programme. I am grateful to all of these institutions for considering my endeavour worthwhile enough for their support.

Abstract

Moral inquiry is often thought of as an individualistic enterprise. This is in large part because morality not only seems to require doing the right thing, but being in touch with moral reality in the right way: to act well, it is necessary to ground our actions on an insight into moral reasons—reasons we can only understand when we inquire for ourselves. In this dissertation, I defend a picture of moral inquiry according to which such inquiry is fundamentally social: we are not alone when it comes to making moral decisions. In particular, I argue that the conception of moral inquiry as an individualistic enterprise wrongly delegitimises a central resource of moral inquiry, moral testimony. To this end, I first argue that moral agreement is of crucial justificatory value in moral inquiry. I then go on to address some worries connected to moral testimony, arguing that reliance on moral testimony is not intrinsically wrong and indeed oftentimes permissible if not required of us. I complement this with a discussion of moral expertise, arguing that the search for moral experts has often been too focused on theoretical knowledge. However, practical experience is a much more reliable way of identifying those we can rely on with respect to a particular moral issue. Lastly, I discuss one consequence this more social picture of moral inquiry has, namely that blame can not only be allocated to a wrongdoing agent, but also to those who guided the agent to that act. This concludes my contribution to the defence of a more social picture of moral inquiry. Overall, however, my aim is not to do away with the importance of being rightly connected to moral reality—it is just that the best connection to moral reality available is often not through ourselves, but through others.

Introduction

Morality is about our interaction with other agents in the world. We are concerned with the effects our choices have on others, how to treat them and how to relate to them. Accordingly, their interests as well as our relationship to them feature prominently when we are deliberating what (not) to do. Morality, to wit, is about a moral *community*—and caring about that community is important.

In spite of this, moral *inquiry* has often been thought of as an individualistic enterprise. How do you come to know what (not) to do? You think, and you think *for yourself*. Other people and their interests do, of course, play a prominent role here, but their role is that of inputs into our own thought process. Other agents are of relevance only insofar as our choices affect them, they are seldom thought of as more than just targets of your actions. To act well, or so many philosophers uphold, you need to base your actions on the right reasons—and to understand those reasons, you need to have thought about them, for yourself. If other people's beliefs come into play at this stage of inquiry, they do so at best only insofar as you think about their beliefs and the reasons they cite in favour of those beliefs, examine them critically, and decide for yourself what's right or wrong.

I have always found some unease with this picture. When listening to others sharing their moral beliefs, I often find myself changing my views on the basis of their diverging beliefs. What is more, I sometimes trust them more or less uncritically—especially when it comes to matters I consider myself to be inexperienced in—and act in ways they think are right; sometimes without understanding exactly why they are. But I trust them to be as conscientious in their judgments as I try to be. Indeed, very often others seem better placed than me to respond to certain moral questions. And so I find myself deferring to others in moral matters on a regular basis. At worst, what follows is an attempt to rationalise my behaviour. At best, it is a series of arguments in favour of extending our understanding of moral inquiry towards a social, rather than an individualistic, enterprise.

This is all very abstract, and it might help to put an example on the table. When thinking about trusting others in moral matters, the following case seems very representative to me:

At a family dinner, I chat to my cousin who is working at a kindergarten. She mentions that the groups in the kindergarten are slightly too big. We talk further about it, but I take her at her word on this matter. If someone were to ask me whether the groups in the kindergarten should be smaller, I will be perfectly happy to say that they should, citing my cousin's testimony in support of my belief.

On the traditional picture of moral inquiry, there is something questionable about my behaviour—and I discuss the reasons to think so in depth throughout this dissertation. For now, however, I should like to note my impression that there seems nothing wrong with my behaviour. In fact, I think such deferral might actually be *required* of me. Why should I, even upon reflection, come to a better conclusion than my cousin, who has received education in pedagogy and has been working in the kindergarten for many years? Of course, the context is doing a lot of work here—and I hope it will become clear why this focus on such context-laden cases is important. Once we place listening to others and trusting them on moral matters into context, I submit, much of what might seem problematic about it disappears.

The picture of moral inquiry I defend in this dissertation, then, is one according to which such inquiry is fundamentally social—we are not alone when it comes to making moral decisions. This picture comes with an emphasis on interacting with others during inquiry, listening to their thoughts, and taking their point of view into account when it comes to making a decision. But this much could be admitted also on the more traditional picture of moral inquiry. What matters, on the traditional picture, is that you are the author—in a substantive sense—of your final assessment, not anyone else. This is where the social picture of moral inquiry most clearly diverges: sometimes, you should be the author of the final assessment in this sense. Sometimes, you should simply *rely* on others and the results of *their* inquiry. It is this latter point that is most contentious and thus the focus of most of my discussion, though that does not mean that it is more important than the former. In short, the social picture of moral inquiry I propose means not just inquiring with other's in mind, but inquiring with others, full stop.

Each of the following chapters considers parts of moral inquiry from different perspectives, and it is possible to read each chapter as a stand-alone paper. Nonetheless, I hope that this introduction can help to outline how they all hang together. Not only do they ultimately have the same aim, the questions some chapters raise are helpfully addressed in others, and the position I defend throughout all of them receives support from each individual chapter.

The first chapter, entitled 'Conciliating to *Avoid* Moral Scepticism', considers the issue of moral *disagreement* and how we should respond to it. Unlike in non-moral disagreements, many philosophers hold that we can retain our moral beliefs even in the face of disagreement. If moral disagreement was reason enough to reduce our confidence in our moral beliefs, and such moral disagreement is widespread, wouldn't we end up with a form of scepticism about morality? In this chapter, I turn this line of thought on its head. I argue that

moral conciliationism, the position advocating for a reduction of confidence in the face of disagreement, *saves* us from epistemological moral scepticism. This chapter simultaneously serves as the motivation for the remainder of the thesis. Once we allow other's moral beliefs to influence ours in this way, it seems like we have opened Pandora's box—we now have to answer *why* it is okay to rely on other's moral beliefs, *whose* moral beliefs we should rely on and *how* we should do so. And so the next chapters address these questions in turn.

In the second chapter, entitled 'The Limits of Pessimism About Moral Testimony', I discuss whether it is okay in the first place to rely on other's moral beliefs, i.e. their moral testimony. Many are pessimistic about this. Here, I argue that influential moral and psychological sources of such pessimism fail. However, I don't think this is because there is nothing to them—in fact, I agree with the pessimist on some of their arguments. But these arguments speak against reliance on moral testimony only under very idealised circumstances, circumstances that we simply don't find ourselves in very often (or ever). As the non-ideal agents we are, I suggest, we often require guidance from others, and so reliance on moral testimony turns out to be perfectly fine.

I continue my exploration of reliance on moral testimony further in the third chapter, entitled 'I think I got it right', in which I present a novel argument in favour of optimism about moral testimony. I argue that it is in principle permissible to rely on one's own moral beliefs formed in the past even if the underlying reasons for the beliefs are no longer occurrent. But there is no (relevant) difference between such *intrapersonal* testimony and the *interpersonal* testimony we are usually concerned with. If so, then reliance on moral testimony should also be permissible. I explore some consequences for both optimism and pessimism about moral testimony following these considerations and argue that the discussion opens up the possibility for a more nuanced optimism about moral testimony.

Assuming that reliance on moral testimony is permissible, one may nonetheless wonder *how* we should take other's testimony or advice into account. The fourth chapter, entitled "Moral Experts: Authorities *and* Advisors", addresses this issue. I argue that we should sometimes treat moral experts as authorities and sometimes as advisors, depending on the *aim* of (a particular) moral inquiry. By default, the aim of moral inquiry is moral understanding and non-experts should therefore treat moral experts as advisors which foster moral understanding. If, however, moral understanding is unavailable or attainable only at high risk, moral knowledge becomes the aim of inquiry and non-experts should treat moral experts as authorities which make moral knowledge easily accessible. Finally, I suggest that the aim of inquiry is in turn determined by the identities we take up in a domain. This explains why the default aim of moral inquiry is moral understanding—we cannot opt out of our identity as moral agents.

In the fifth chapter, entitled 'Moral Expertise and Moral Philosophy' I continue the discussion of moral expertise and address the question *whom* we should listen to when we seek out other people's moral testimony. I argue that the common conception of moral experts as focusing on theoretical knowledge of morality misses its target. Instead, we should discuss moral expertise more

fine-grained, with respect to the several subdomains of morality. Shifting the focus in this way suggests that practical experts—activists, those (negatively) affected, etc.—are much more likely to provide reliable moral testimony than, as is often held, moral philosophers.

I hope that chapters 2 to 5 in conjunction provide both a motivation to think that reliance on moral testimony is often permissible (or even required) as well as do away with some of the initial objections and worries we might have about this new picture of moral inquiry. Moral inquiry, these chapters suggest, is more social than we used to think because listening to others and trusting them on their judgments is desirable for its consequences of enabling us to act well, but also in itself. In the final chapters of the thesis, I turn to some consequences this picture of moral inquiry has.

This sixth and final chapter, entitled ‘Moral Testimony and Blame’ considers the question who gets blamed when we rely on other people’s moral testimony and things go wrong. On the traditional conception of moral inquiry, responsibility for acts is assigned to the acting agent themselves. In considering whether an agent is blameworthy, we usually draw on their reasons for acting. But what if their reasons for action are, following the nature of my project here, socially acquired reasons? I argue that both the wrongdoing agent and the testifier can be blameworthy—and can, in particular, legitimately be blamed by the wronged agent. I draw attention to the fact that taking both agents seriously as actors entails conceiving of both as responsible agents—in their deferring and in their testifying. This helps us to circumvent worries about shirking responsibility when it comes to reliance on moral testimony: relying on someone else’s advice does not get you off the hook. At the same time, it also has some consequences for testifiers: sharing your moral beliefs, I suggest, means taking some responsibility for the actions taken on the basis of those beliefs—even if their enactment is outside your control. Relying on and sharing moral testimony, then, is as serious a part of moral discourse and practice as any other form of deliberation. I suggest that pessimism about moral testimony cannot do justice to this result, while optimism about moral testimony can.

Following these six chapters, I have included some concluding remarks, reflecting on the findings throughout this dissertation, and pointing out the most important connections. But more than that, I hope that the conclusion serves as a sort of outlook. There are important questions that I did not have the time or space to address in this dissertation, and so there is as much left to do as there is done here.

Chapter 1

Conciliating to *Avoid* Moral Scepticism

A common worry about moral conciliationism is that it entails at best uncertainty about many of our moral beliefs and at worst epistemological moral scepticism. Against this worry, I argue that moral conciliationism *saves* us from epistemological moral scepticism and enables us to be confident in many of our moral beliefs. First, I show that only taking disagreements seriously as a threat to our beliefs allows us to utilise agreements in support of our beliefs (call this symmetry). Next, I argue that utilising moral agreements as an epistemic resource allows moral conciliationism to resist the potentially worrisome reduction in confidence of our moral beliefs. Taking the relevance of moral agreement into account, I argue that it is *anti*-conciliationism that must meet the challenge of epistemological moral scepticism. For this, I suggest that moral inquiry is best understood as a collective endeavour. If so, then agreement on our moral judgments is required to justify the confidence we have in many of our moral beliefs. However, by symmetry, this appeal is possible only if one accepts the conciliatory attitude towards disagreements. Hence, *accepting*, rather than rejecting moral conciliationism, is the way out of moral scepticism.*

*A version of this chapter has been awarded with the 2022–23 Robert Papazian Prize for Essays in Ethics and Political Philosophy and has subsequently been published in the *International Journal of Philosophical Studies* (cf. Küspert (2023)).

1.1 Introduction

I believe that meat consumption is morally impermissible. At the same time, I know that a significant number of ethically capable people disagree: they believe that meat consumption is morally permissible.¹ Should this undermine the confidence I have in my belief?

The standard attitude towards cases of moral disagreement is that moral steadfastness—retaining one’s moral beliefs in the face of disagreement—is epistemically permissible or perhaps even required (cf. Elga (2007), Setiya (2010) and Setiya (2012)). Such steadfastness is opposed to moral conciliationism. According to moral conciliationism, the correct response to moral disagreement is to conciliate: acknowledging that a significant number of ethically capable people disagree with me about the moral status of meat consumption should reduce my confidence in the respective belief for I cannot discount my interlocutor’s belief solely on the basis of the present disagreement. Many worry that moral conciliationism thus entails an epistemological kind of moral scepticism. Since moral disagreement is widespread and persistent, the conciliatory attitude seems to demand that we significantly reduce our confidence in many of our moral beliefs thus leading to said scepticism (cf. e.g. Feldman (2006a), Feldman (2006b), Christensen (2007) and Kornblith (2010)). If such scepticism is to be avoided, we should better reject moral conciliationism.

Vavova (2014) argues that the charge of a wholesale scepticism isn’t quite warranted. She holds that conciliationism is committed only to agnosticism about *controversial* moral beliefs. However, many of our more basic moral beliefs are met with so much agreement that conciliationism is not committed to a reduction in confidence in these beliefs, thus admitting at best some sort of ‘local’ moral scepticism.

In this chapter, I begin by investigating the theoretical foundations of this kind of defence of conciliationism, arguing that it requires a claim about the *symmetry* between the epistemic significance of disagreement and agreements. Only taking disagreements seriously as a threat to our beliefs allows us to utilise agreements in support of our beliefs, thus conferring the certainty of the more basic moral beliefs that are met with a lot of agreement. Having established this symmetry claim, I argue that *only* moral conciliationism can therefore draw on the epistemic significance of moral agreements. The anti-conciliationist, by contrast, cannot draw on agreements in the same way to vindicate their moral beliefs. Thus, if there is any threat of moral scepticism, it is—perhaps surprisingly—to be found on the *anti*-conciliationist side. I suggest that moral inquiry is best understood as a collective endeavour. If so, appealing to agreements on our moral judgments is required to justify the confidence we have in our moral beliefs. However, by symmetry, this appeal is possible only if one

¹A 2010 study finds that there are about one and a half billion vegetarians worldwide, only 75 million of which are vegetarians by choice (cf. Leahy et al. (2010)). Arguably, a significant number of non-vegetarians by choice believe that meat consumption is morally permissible, many of whom are likely to be ethically capable.

accepts the conciliatory attitude. Hence, *accepting*, rather than rejecting moral conciliationism, is the way out of moral scepticism.

My defence of moral conciliationism here thus rests on a bigger conceptual issue—understanding the positions in the literature around disagreement not just as positions about disagreement, but as positions about the value of *testimony* in general, regardless of whether the testimony happens to be in agreement or disagreement with prior beliefs. While I do not explore this underlying issue in more detail here, I hope that it provides sufficient motivation to explore the debate around disagreement in this more open-ended inquiry regarding testimony in general.

Finally, let me be explicit that my concern here is solely epistemological. Sometimes, moral disagreement is also invoked as an argument for *metaphysical* moral scepticism. The persistence of moral disagreement is thought to be best explained by appealing to different ways of life rather than disagreement about some moral reality (cf. Mackie (1977), 36). For the purposes of this chapter, I set such worries aside. After all, the anti-conciliationist is worried that moral conciliationism might lead to *epistemological* moral scepticism. Accordingly, metaphysical moral scepticism most certainly is a non-starter for them.

1.2 Symmetry

An initial worry about moral conciliationism is that it might entail a wholesale epistemological moral scepticism. If moral disagreement is widespread and pervasive, affecting all or at least most of our moral beliefs, it seems that it might undercut our justification for holding moral beliefs entirely. In her 2014 paper, Vavova argues convincingly that such wholesale scepticism is not an outcome of any plausible version of moral conciliationism. She argues that there are many moral beliefs, especially our most basic moral beliefs, for which disagreements are rare and agreements much more prevalent (cf. Vavova (2014), 302).

I am sympathetic to this kind of approach for defusing the initial worry about a far-reaching scepticism. However, I think that the argumentative move Vavova utilises merits further attention. Note that she is not just claiming that disagreements are rare for many of our moral beliefs, she is also invoking *agreement* as a counteracting force. The best explanation to ground this argumentative move, I submit, is an appeal to a symmetry between the epistemic significance we assign to disagreements and the epistemic significance we assign to agreements. The symmetry claim then holds that whatever the epistemic significance of disagreement, the same goes for agreement and vice versa.² I

²Epistemic significance comes in many ways. For my purposes here, I am interested in a notion of epistemic significance that has an effect on your *justification* for the (dis)agreed upon belief. The symmetry claim thus holds that agreements affect your justification to the extent that disagreements affect it and vice versa. In the following, I often use the qualifier ‘substantial’ to mark the difference between moral conciliationism and the anti-conciliationist. Roughly, the epistemic significance of (dis)agreements is substantial if (dis)agreements usually have noticeable effects on the justification of our beliefs—it is this noticeable effect that gets

thus suggest understanding Vavova as arguing that any reduction in justification following from rare instances of disagreement about many moral beliefs is outweighed by the much more frequent agreement about these beliefs. Appealing to symmetry, the conciliationist thus has a resource to argue that there is no need to reduce our confidence in all of our moral beliefs given the significant amount of agreement on many of them.

Why think that symmetry holds? Both disagreements and agreements are concerned with the epistemic significance of an interlocutor's belief regarding a proposition you have already entertained. Thus, the setup of the two cases is symmetric. In both cases, you receive testimony regarding a claim you have already considered. Whether this testimony happens to align with your beliefs should be irrelevant for the epistemic significance you ascribe to it, because the level of trust in a particular piece of testimony depends on the antecedent trust in the testimony's source, in this case your interlocutor (cf. Fricker (2006), 600). When considering peer (dis)agreements, this claim is particularly plausible—after all, taking someone to be your peer just means that you trust them as a source of testimony in general, to the point that you think they are as likely as you to get things correct. Only accepting symmetry allows you to take your peers *themselves* seriously as a source of evidence.

Though the symmetry claim has, to the best of my knowledge, not been explicated in the literature so far, I take it that it underlies Vavova's argument and is in fact prevalent even in the anti-conciliationist camp. For a very explicit example, consider Fritz (2018)' argument *against* conciliationism: because a wide backdrop of moral agreement is a good reason to trust strangers in moral matters, he argues, moral conciliationism is committed to significant reductions in confidence in the face of moral disagreement, thus leading to objectionable spinelessness (cf. Fritz (2018), 166). Such dependence of the epistemic significance of disagreements on the epistemic significance of agreements requires the assumption of something in the spirit of the symmetry claim.

In this section, I consider various objections to the symmetry claim. Given that the symmetry claim has not been explicated in the literature so far, the objections I discuss draw on discussions which are not originally intended as objections to symmetry. I begin with two objections claiming that disagreements have substantial epistemic significance while agreements lack it. After dealing with these, I discuss two objections that challenge symmetry the other way around. These aim to undermine the epistemic significance of disagreements while leaving the epistemic significance of agreements intact.

Since the symmetry claim has a broader scope than just *moral* agreements and disagreements, the examples I draw on in this section are not restricted to the moral domain. The symmetry of *moral* agreement and disagreement is just a special case of a more general symmetry claim about agreement and disagreement.³

conciliationism out of the charge of scepticism. In contrast, for the anti-conciliationist, the justification of our beliefs is usually not noticeably affected by (dis)agreements.

³Of course, one may hold that symmetry holds in general, but not in the moral domain. The burden of proof, however, is on those upholding this position—they would have to provide

1.2.1 Belief

Let us begin by considering potential problems for symmetry alleging that sometimes agreements lack epistemic significance while disagreements do not. Is the mere fact that you already believe something a potential symmetry breaker when it comes to agreement and disagreement on that belief? Roughly speaking, if you believe a proposition, any agreement on it might seem epistemically vacuous—responding to the agreement you simply continue to believe the proposition. In contrast, when encountering a disagreement, your doxastic states may change: you can hold on to the belief but question it, you can give the belief up, etc.

One way to substantiate the present intuition is by drawing on the ‘default and challenge’ model of justification. Following this model, you are entitled to believe some proposition until a justified challenge to the proposition comes up (cf. e.g. Williams (2001), 149). This model has most plausibility when applied to basic sources of knowledge such as perception or even intuitions. If so, then on many views according to which intuitions play a central role in moral epistemology, it also has application in the moral realm. Huemer’s phenomenal conservatism, cited in support of his intuitionism, is one such application of the default and challenge model (cf. Huemer (2005), 99). Following phenomenal conservatism, one is entitled to hold a moral belief based on one’s intuition as long as the belief remains unchallenged. If one considers moral disagreements to constitute such a challenge, then symmetry fails. While moral disagreements have epistemic significance in that they challenge your beliefs, the same does not seem to hold for moral agreements. After all, or so the argument goes, one’s justification for the initial belief stems from one’s intuition, not potential agreement on the belief.

To see whether this objection actually threatens symmetry, let us consider the reasons for why decreasing one’s confidence in a belief seems justified in such cases of disagreement more closely. Suppose I am very confident in my belief that modus ponens holds. Since I take myself to be in a good position to evaluate the issue (otherwise my confidence would not be as high), I take myself to be capable with respect to this issue. Naturally, then, I also expect other capable people to come to the same conclusion. Following up on such an agreement with a capable agent, nothing much about my evidential status changes. After all, I expected capable people to agree with me, I took my interlocutor to be capable and they agreed with me. Thus, the agreement at hand confirms the hypothesis I already had a high confidence in—it seems to lack substantial epistemic significance. In contrast, in cases of disagreement, my evidential status changes quite drastically. Again, I believed that capable people would agree with me, I took my interlocutor to be capable but nonetheless they ended up disagreeing with me. Hence, some of my prior hypotheses must be false. Either it is false that capable people come to my conclusion or my assumption that my interlocutor is capable is false (of course, the conciliationist should take the

a (non-circular) argument as to why the moral domain should not be subject to symmetry when other domains are.

first of these options). Either way, the disagreement has substantial epistemic significance.

It should become clear now, however, that the reason for the differing epistemic significance of agreement and disagreement is not, as initially suggested, due to the very nature of agreement and disagreement. Rather, it is due to my prior *expectations* about the likelihood of agreements and disagreements. If I come to believe a proposition, I should also form further beliefs, one being that capable people will agree with me. Hence, I have already factored in that capable people will agree with me when forming my initial belief. Thus, if someone whom I take to be capable agrees with me, my evidential basis changes little. After all, I *expected* my interlocutor to agree with me. By contrast, if said interlocutor disagrees with me, my evidential basis changes significantly—it is for this very reason that my belief is challenged.

However, the difference in epistemic significance is not inherent to agreements or disagreements. Had I expected a *disagreement*, its epistemic significance would have been minor while an agreement would have had more epistemic significance. Suppose, for example, that you have a discussion about a hot political issue with someone from the other side of the aisle. You fully expect the two of you to disagree since you disagreed with almost all claims of your interlocutor on the issue at hand. However, to your surprise, you find them agreeing with you. Given your prior expectations, you must now either modify your estimate of their competence on this issue or begin questioning your own belief. Long story short, the agreement has substantial epistemic significance because you did not expect its occurrence.⁴

My expectations about the likelihood for an agreement or a disagreement thus affect their respective epistemic significance. If I expect them, their epistemic significance is low, if I do not, their epistemic significance is high. This suggests that in neutral cases, where I have no expectation about whether I will encounter an agreement or a disagreement, both agreements and disagreements are of equal epistemic significance—and this is just the most basic version of the symmetry claim we are out to defend.

Given these considerations, however, we can now formulate the symmetry claim somewhat more rigorously. The claim is not simply that all agreements and disagreements are of equal epistemic significance. Rather, there is nothing about agreements and disagreements *themselves* that gives one or the other more epistemic significance. Evidently, external factors, such as expectations about the likelihood for an agreement or a disagreement might have an impact on their epistemic significance. However, the mere fact of agreement or disagreement does not entail a difference in epistemic significance. The symmetry claim is thus to be understood as an ‘other things equal’ claim. Other things equal, there is no difference in the epistemic significance of agreements and disagreements.

⁴Psychological experiments confirm this extra-attentiveness to information that contradicts our expectations whether in agreement or disagreement (cf. Sperber and Mercier (2017), 210 for an overview of some such experiments).

1.2.2 Reasons for Belief

So far, we have focused on the mere fact of agreement or disagreement. However, instead of drawing on these as the symmetry breaker, one might propose drawing on the nature of the *reasons* underlying the beliefs leading to agreements and disagreement. Let us distinguish between reasons for and reasons against a belief (cf. Snedegar (2018)). If someone disagrees with you, it seems that you will usually learn of a new reason against your belief—the reason that led them to come to the opposite belief. In contrast, agreements can occur also when the agreeing party relies on the same reason to justify their belief. But if so, or so the thought goes, they do not provide you with a new reason for your belief. Thus, symmetry fails.⁵

Suppose that my belief in the impermissibility of meat consumption is due to a further belief that abstaining from meat consumption minimises our impact on ecosystems (cf. Fox (2000) for this position). Imagine I encounter an agent whom I take to be ethically capable who agrees with me that meat consumption is morally impermissible. When inquiring about their reasons for holding this belief, they cite exactly the same reason I relied on to begin with. Evidently, I do not have two reasons going in favour of my belief now—the reasons we cite are identical, in similar cases they may at least overlap (cf. Nair (2016) and Maguire and Snedegar (2020)). At any rate, the epistemic significance of the agreement is significantly reduced. But the same does not hold for disagreements, or so the argument goes. For suppose the agent I encounter disagrees with me, arguing that we have a fundamental right to choose what we consume, including meat in particular. If I take this reason on face value, I now have a reason for and a reason against my belief. The disagreement is of epistemic significance—it provides me with a new reason while the agreement does not. And if the epistemic significance of agreement is thus undermined while the epistemic significance of disagreement is not, then symmetry fails.

While I think that the characterisations of these cases are correct so far, I do not think that symmetry fails as a result. Rather, the examples show that it is worth differentiating different kinds of agreements and disagreements respectively. I show this in the following by providing examples of disagreements that have equal epistemic significance as the agreement considered above (for lack of introducing new reasons) as well as examples of agreements that are just as epistemically significant as the disagreement considered above (because they introduce new reasons). Let me begin with the latter, more obvious point.

It is of course possible to agree with someone on a proposition, even though they provide different reasons for believing the proposition than you. Suppose the concurring agent I encounter argues that meat consumption requires us to treat animals as mere means which is morally impermissible (cf. Korsgaard

⁵In offering this explanation as to why symmetry might fail, it may seem that we are leaving the territory of discussing the epistemic significance of agreements and disagreements *themselves*, discussing only the underlying reasons instead. However, if we could show that disagreements *in general* provide new reasons while agreements fail to do so, we have nonetheless shown something about the nature of agreements and disagreements.

(2012) for such a position). Assuming that I did not consider this reason before, I am now provided with a new reason for my belief. Hence, these kinds of agreements might well have substantial epistemic significance.

There are, then, at least two kinds of agreement. First, there are agreements relying on the same source—the epistemic significance of these agreements may be significantly weakened since they merely provide you with higher order evidence that you appreciated the reason in question correctly. Second, there are agreements drawing on distinct sources—the epistemic significance of these agreements is not under threat from the present argument since you are provided with a new reason for.⁶ Note that this distinction also has a desirable normative upshot: on this account, typical echo chambers do not provide you with a substantial justification for your beliefs because only agreements providing new reasons do so.

So, there are two kinds of agreement, one of which plausibly has substantial epistemic significance, one with weakened epistemic significance. Still, if all kinds of disagreements carry equal epistemic significance, symmetry nonetheless fails.

Though a disagreement cannot, as it were, *repeat* one’s initial reason as in the first kind of agreement, it can still be *about* the initial reason. Suppose for example that my disagreeing interlocutor, after being provided with my reasoning relying on the conversation and restoration of ecosystems, argues that they have considered exactly this reason, but took it to provide no support for the belief that meat consumption is morally permissible. My conservative approach to ecosystems, they take it, rests on some kind of a naturalistic fallacy. The connection I drew between the reason at hand is dismissed by my interlocutor.⁷ Such disagreements invite you to reconsider whether or not you have appreciated the respective reason correctly. The disagreement indicates that you may have gone wrong at some point.

We are now in a position to fully appreciate the potential asymmetry. If you believe some proposition, you usually have a reason for doing so. Agreements can thus take two forms, they either repeat your initial reason or they provide you with a new reason to believe the proposition. The former lacks substantial epistemic significance while the latter has it. In contrast, disagreements always seem to carry substantial epistemic significance. Either they give you higher-order evidence that you may have failed to appreciate the initial reason correctly or they give you a new reason against the proposition in question.

The threat of asymmetry thus comes down to whether there is a difference in epistemic significance between agreements repeating your initial reason and disagreements threatening your interpretation of your initial reason. However,

⁶In a footnote, Vavova alludes to this kind of agreement and notes that their epistemic significance derives from the “independent support for our judgments about the [beliefs]” (Vavova (2014), 332).

⁷In other cases, the reason in question might even be taken as a reason against by my interlocutor. In the above example, they might for example believe that humans are in fact *meant* to shape and modify ecosystems in a way that suits their needs best. My reason for would be their reason against.

refining the worry in this way undermines its initial force. For the initial charge was that some agreements ‘give you nothing new’ while disagreements always ‘take something away’. Comparing the cases of (dis)agreement which either support or threaten your appreciation of a particular reason does not present this strong intuitive difference. Both the agreement and the disagreement affect the support for the reason you already have considered. While the agreement provides you with no new reason for, neither does the disagreement provide you with a new reason against. Rather, the agreement supports the connection you drew between reason and proposition in question while the disagreement threatens it. Both such agreements and disagreements are higher-order evidence about your already present reasons. While the agreement does not add a new reason, neither does the disagreement. If at all, it threatens your interpretation of the already present reason, but the reverse is true for the agreement—it supports your interpretation of the reason. In fact, these are exactly the cases the traditional peer disagreement literature tends to focus on.

Similar remarks apply to (dis)agreements about how to correctly weigh the reasons for and against a proposition. The degree to which this higher-order evidence increases or decreases your confidence in the (dis)agreed upon weighing of reasons may differ categorically depending on how confident you are in its connection to the supported proposition. Crucially, however, the epistemic significance does not differ for cases of agreement and disagreement. The intuitive difference between the two cases is best explained by way of your expectations, as I argued in the previous section. This concludes my defence of symmetry in the first direction: if disagreements have substantial epistemic significance, so do agreements.

1.2.3 Permissivism

Let us now turn to objections challenging symmetry in the other direction. These objections maintain that disagreements can lack substantial epistemic significance even if agreements do have substantial epistemic significance.

A major motivation for conciliationism stems from the conviction that for any set of evidence, there is a uniquely rational doxastic attitude for any proposition. This conviction is generally referred to as the uniqueness claim (cf. e.g. White (2005) and Feldman (2006b)). Roughly, given uniqueness, disagreement on the ground of shared evidence entails that at least one party to the disagreement has failed to form that unique rational doxastic attitude. However, though widespread, the acceptance of uniqueness is by no means trivial. The contender is permissivism, which denies uniqueness. According to permissivism, at least sometimes, given a set of evidence, there is more than one rational doxastic attitude for a proposition (cf. e.g. Schoenfield (2014) for a defence of permissivism).

Permissivism might undermine symmetry as follows. According to permissivism, disagreements are not necessarily an indicator for a mistake in rationality on behalf of a party to the disagreement—after all, there may be multiple rational doxastic attitudes towards the proposition in question. Hence, according to permissivism, there seem to be disagreements with no or only weak epistemic

significance. However, there is, at least on the face of it, no reason to assume that permissivism likewise undermines the epistemic significance of agreements. After all, permissivism is usually not understood as an ‘anything goes’ position. Hence, even if permissivism is true, agreements are epistemically significant. They provide some (defeasible) evidence that one has formed one of the available rational doxastic attitudes.

Responding to this objection will also allow me to refine my understanding of epistemic significance further. On a naïve understanding, new evidence is epistemically significant if and only if it either weakens or strengthens the justification for the belief you hold. However, while the disagreement does not weaken or strengthen the justification of your current belief on the permissivist account, the disagreement still provides you with new information. After all, you receive significant information about the set of evidence: not only is your own belief a rational doxastic attitude towards it, also the belief of your interlocutor may be such.⁸ Therefore, the disagreement still carries epistemic significance in the sense that it affects the justification you have available for that belief. While it does not undermine your current belief, it affects your justification for other rational doxastic states towards the proposition. Your current doxastic state towards the proposition is justified and the doxastic state of your interlocutor may be as well.

Interestingly, not even accepting an ‘anything goes’ version of permissivism threatens symmetry (Christensen (2016), 587–588 briefly outlines such a position). For while accepting such an extremely permissive position might entail that disagreements really are epistemically insignificant (since anything goes, learning that some particular belief goes as well is not new information), it likewise entails that agreements are epistemically insignificant. Learning that there is agreement on your belief carries no new information about the justification of said belief.

Thus, permissivism is consistent with symmetry. This is a desirable result for two reasons. The first is obvious: accepting permissivism does not threaten symmetry. The second is more interesting. If symmetry holds both if one accepts uniqueness and if one accepts permissivism, then accepting symmetry does not entail any further commitment in this debate. While accepting symmetry evidently entails some commitments (after all, it is a non-trivial claim), it is non-committal in the debate around uniqueness and thus does not require prior acceptance of one of the major motivations for conciliationism.

1.2.4 Extreme Disagreement

There is one class of disagreements that even committed conciliationists usually consider to be void of epistemic significance: extreme disagreements (cf. Christensen (2007)). One way to undermine the symmetry claim, then, is to

⁸This does not hold for just any interlocutor. Only the disagreement of epistemically capable agents provides you with this information. However, this does not undermine symmetry. Likewise, only the *agreement* of epistemically capable agents affects your justification in a positive way.

argue that a significant number of *moral* disagreements are such extreme disagreements. If so, moral disagreements lack epistemic significance while the epistemic significance of moral agreements can be retained and, accordingly, symmetry fails.

To get an intuitive grip on extreme disagreements, consider the following case:

Suppose that five of us go out to dinner. It's time to pay the check, so the question we're interested in is how much we each owe. [...] I do the math in my head and become highly confident that our shares are \$43 each. Meanwhile, my friend does the math in her head and becomes highly confident that our shares are \$450 each. (Christensen (2007), 199, my emphasis)

Extreme disagreements, then, are such that the response of your interlocutor is, for some reason or other, out of the realm of the options you would seriously consider. In the above example, the share of each person cannot possibly exceed the total bill, \$450 is thus a non-starter. You may think that the same is true in many cases of moral disagreements. Consider, for example, the belief that kicking puppies for fun is morally permissible. Adopting this position is a similar non-starter as the \$450 in the above example. A disagreement of this kind is thus intuitively void of epistemic significance.

Note first that, since there are difficult moral questions in which several outcomes are at least on the ballot, not *all* moral disagreements can be extreme. Drawing on extreme disagreements to undermine symmetry will therefore work only for a limited class of disagreements (as much can readily be admitted by the proponent of this symmetry breaker). However, even this restriction is not sufficient to deny symmetry. For if extreme moral disagreements are only those in which your interlocutor is, from your point of view, fundamentally mistaken about a moral issue, then you have little reason to think of them as ethically capable—after all, they are, by your own lights, *fundamentally* mistaken (cf. Vavova (2014) for this line of argument).⁹ To generalise, the more extreme the disagreement, the less reason you have to think of your interlocutor as ethically capable. However, moral conciliationism was never meant to apply to all moral disagreements but only to disagreements with ethically capable agents. Ironically enough, then, the more extreme the disagreement, the less adjustment to your beliefs conciliationism typically recommends.

The reason for the lack of the epistemic significance in these cases is not the extremeness of the disagreement, but the ethical incapability of your interlocutor. Hence, the underlying rationale to neglect the epistemic significance of

⁹Many commentators think that such reasoning is in violation of the independence criterion. But if so, then *extreme* disagreements deserve no special treatment and the attempt to draw on them to undermine symmetry fails from the outset. An alternative proposal as to why extreme disagreements are epistemically insignificant comes from Lackey and draws on the prior confidence one has in the belief (cf. Lackey (2010)). Again, extreme disagreements are rendered epistemically insignificant. However, as I argued earlier, the same goes for agreements on beliefs in which you were highly confident to begin with.

extreme disagreements is to neglect the epistemic significance of disagreements with ethically incapable agents. Plausibly though, the same idea applies to agreements as well: the less ethically capable you take your interlocutor to be, the less epistemic significance your agreement with them has.

The relevant contrast to standard extreme moral disagreement is thus not just any moral agreement but a moral agreement with an ethically incapable agent. Notably, however, symmetry *does* apply here: the epistemic significance of both disagreements and agreements decreases with the ethical capability of the interlocutor. Hence, drawing on extreme disagreements to reject symmetry does not work. Still, this discussion clarifies once more that symmetry does not entail that *all* agreements and disagreements are of equal epistemic significance. Rather, the epistemic significance of both agreements and disagreements depends (and depends to the same degree) among other things on the ethical capability of your interlocutor.

For each objection, then, the underlying intuition can be traced back to a source which, if properly investigated, fails to support the charge against symmetry. Thus, symmetry holds: other things equal, whatever the epistemic significance of disagreement, the same goes for agreement and vice versa.

If moral conciliationism holds, then moral disagreements carry epistemic significance—but, given symmetry, moral conciliationism also entails that moral agreements carry epistemic significance. This is why arguments such as the one Vavova presents in defence of moral conciliationism work. Similarly, by symmetry, if the anti-conciliationist claims that moral disagreements lack epistemic significance, they are also committed to the claim that moral agreements lack epistemic significance. This observation will be the first building block of the positive argument against anti-conciliationism to be explored in the next section.

1.3 Moral Agreement

By investigating the symmetry claim underlying the defence of moral conciliationism, we have seen that moral agreements might play a more important role than they have typically been assigned in the debates around peer disagreement. In this section, I suggest that utilising moral agreement as an epistemic resource can save us from epistemological moral scepticism and that failing to utilise it in this way may, under certain assumptions, lead us to said scepticism. This discussion becomes possible because of the symmetry claim: I have shown that it is not necessary to discuss the epistemic significance of moral disagreements directly—it is equally promising to discuss the epistemic significance of moral agreements.¹⁰

¹⁰The focus on agreements also brings up an interesting implication for one particular version of anti-conciliationism, the so-called total evidence view (cf. Kelly (2010)). On this account, one ought to take into account both one's evidence prior to a disagreement and the fact of disagreement, thus sometimes allowing to discount one's interlocutor *on the basis of* the present disagreement (i.e. denying the independence criterion conciliationism subscribes to). Focusing on agreements, denying independence has the surprising consequence that we can sometimes 'count' an interlocutor's opinion on an issue solely on the basis of our

It may seem that establishing symmetry is of little value for the discussion around moral conciliationism and epistemological moral scepticism. After all, arguing for moral conciliationism initially required an argument to the extent that moral disagreements have substantial epistemic significance. All the symmetry claim has established is that moral conciliationism can also be established by showing that moral *agreements* have substantial epistemic significance. However, this difference is not to be neglected. The anti-conciliationist attitude towards cases of moral disagreement is to deny their normative force. In contrast, as I hinted at when first introducing the symmetry claim, anti-conciliationists regularly point out the prevalence of moral agreements, drawing on their epistemic significance in order to show that moral conciliationism cannot simply disregard ‘inconvenient’ moral disagreements. Let us consider again Fritz’s (2018) note on why moral conciliationism entails a dispute-independent reason to trust even total strangers about morality, this time quoted at length:

Although the practical importance of reaching moral consensus tends to draw our attention towards areas of dispute, most people probably *share* the *vast majority* of their moral beliefs. [...] When we note that other people generally seem to be right about a wide range of moral questions, we gain reason to believe of any given stranger that [they] will be right about a wide range of moral questions. (Fritz (2018), 108, my emphasis)

Fritz draws on this result to show that moral conciliationism cannot simply disregard moral disagreements in the way anti-conciliationism supposedly can. Fritz’s appeal to moral agreement, however, also shows that even anti-conciliationists tend to assign epistemic significance to moral agreements.

Therefore, the route via the epistemic significance of moral agreements is preferable to the route via the epistemic significance of moral disagreements. For in the case of moral agreements, we find at least some common ground when it comes to their epistemic significance. Hence, we need not refute the anti-conciliationist at the very outset of the project.

1.3.1 Moral Conciliationism does not entail Scepticism

Recall the initial charge of the anti-conciliationist. Supposedly, moral conciliationism leads to epistemological moral scepticism because moral conciliationism assigns substantial epistemic significance to moral disagreements and such moral disagreements are widespread. However, as Vavova notes, this argument succeeds only if moral disagreements are so widespread that they affect even our basic moral beliefs (cf. Vavova (2014), 302).

To see why the charge against moral conciliationism fails, let us first investigate the effects of symmetry on moral conciliationism. According to symmetry, the epistemic significance of disagreements and agreements are, other things equal, the same. If so, then a proper characterisation of moral conciliationism entails not only that moral disagreements are of substantial epistemic

agreement with them on that same issue. This seems worrisome given that it might enable us to ‘overproduce’ peers whose agreement we would then take to support our initial belief.

significance, but that moral agreements are likewise of substantial epistemic significance.

If, as the anti-conciliationist holds, the epistemic significance of moral disagreements is such that, following moral conciliationism, it can lead to epistemological moral scepticism, then the reverse should hold for moral agreements. Following moral conciliationism, moral agreements are ascribed such epistemic significance that they can lead to stable moral beliefs, i.e. avoiding epistemological moral scepticism.

Thus, for the charge against moral conciliationism to be successful, the anti-conciliationist needs to argue not just that moral disagreement is widespread but also that moral conciliationism cannot draw on moral agreements to avoid scepticism. However, as we have seen in establishing symmetry, it is not clear why moral conciliationism should be unable to draw on moral agreements. After all, in assigning epistemic significance to moral disagreements, moral conciliationism can (and should) similarly assign epistemic significance to moral agreements.

While moral conciliationism does advise a sceptical attitude about moral beliefs that are subject to persistent and widespread disagreement (this is the initial charge of the anti-conciliationist), the sceptical attitude is not principled. If the disagreement about the moral belief in question resolves, it is possible to draw on the new-found agreement as a justifying factor in favour of believing it. I agree here with Shafer-Landau (2003) that there seems to be at least some moral progress in society, and—like him—I do not think that a resort to future agreement is naïve (cf. Shafer-Landau (2003), 16–18). While accepting moral conciliationism may mean accepting uncertainty about some moral beliefs (given that they are subject to disagreement), it also offers a way out of the uncertainty.¹¹

Appealing to the difference between transitional and terminal attitudes is instructive to clarify this point. Transitional attitudes are part of the *process* of reasoning while terminal attitudes are *conclusions* of reasoning (cf. Staffel (2019)). Moral conciliationism, I suggest, is best understood as advising suspending belief *for the time being*. Yet, the suspension of judgment is transitional, not terminal—the moral conciliationist offers a way out of the suspension of judgment. Because of symmetry, moral conciliationism has a good explanation for the confidence we have in many of our moral beliefs—the substantial epistemic significance of moral agreement.¹²

¹¹These considerations strongly suggest that there is some sense in which numbers matter when it comes to agreements and disagreements. For an overview of the potentially connected worries, cf. Frances and Matheson (2019), section 6.

¹²In a later paper, Staffel suggests her own view on higher-order evidence that one may have reasoned incorrectly, the *unmooring* view. Similar to my suggestion here, she suggests that higher-order evidence that an agent may have reasoned incorrectly (e.g. peer disagreement) entails that one's attitude towards the belief is no longer justified as a terminal attitude. It might, however, be justified as a transitional attitude. Nonetheless, and drawing on an example of peer disagreement, she agrees that even this might be undermined sometimes such that conciliating also on the level of transitional attitudes is appropriate (cf. Staffel (2021), 249).

And such moral agreement is indeed abundant. Note first that a certain amount of moral agreement seems to be required for the proper functioning of a society. There are, as e.g. Vavova observes, arguably many basic moral beliefs on which we converge, “that pain is bad, that we should take care of our children, etc.” (Vavova (2014), 302). Next, as indicated in the earlier passage I quoted from Fritz (2018), even though we tend to *focus* on moral disagreement, both in the political realm as well as in the philosophical literature, this should not lead one to conclude that the amount of moral agreement is substantially limited. In fact, it may be that the worries we have about moral disagreement show that we are used to a substantial backdrop of moral agreement.

Even drawing on different moral theories—e.g. utilitarianism, virtue ethics or Kantianism—that are, or so it seems, in clear disagreement, does not help much. While there are deep disagreements about the justification of our moral beliefs, the disagreements rarely translate to first-order disagreements:

For example, both Mill and Kant can agree that we shouldn’t kick puppies. Mill would think this for the standard utilitarian reasons: puppies are sentient creatures, pain is bad, etc. Kant seems to think we shouldn’t kick puppies because it corrupts our moral character and increases the likelihood of us kicking morally relevant creatures like humans. But still, he does think we shouldn’t kick puppies. (Vavova (2014), 323)

Again, it may be that we tend to focus on the disagreed upon first-order questions (such as trolley problems and the like), but this is mostly to bring out intuitions about the differences in justification of the theories. The actual first-order disagreement between (plausible) moral theories is limited.

1.3.2 Moral Inquiry as a Collective Enterprise

So far, I have explicated an underlying assumption employed in the defence of moral conciliationism (symmetry) and then argued that symmetry ensures that epistemological moral scepticism does not follow from moral conciliationism. The new focus on *agreement* rather than disagreement, now allows us to turn this defence into a positive argument against the contender of moral conciliationism. After all, symmetry shows that the denial of the epistemic significance of disagreement goes hand in hand with a denial of the epistemic significance of agreement.

What, then, if we reject the epistemic significance of moral agreements? After all, if symmetry holds, this rejection entails that moral disagreements are likewise void of epistemic significance—so perhaps the anti-conciliationist can nonetheless uphold many of their moral beliefs (while they cannot draw on moral agreements, they seemingly do not ‘need’ to, given that there is no reduction in confidence due to moral disagreements in the first place).¹³

¹³In the literature on moral testimony, some hold that moral testimony fails as a reliable source of evidence, either because we cannot identify reliable testifiers in the moral domain (cf. McGrath (2009)) or because of the high stakes in moral matters (cf. Hopkins (2007)). However, most worries about moral testimony arise not due to such *unavailability* concerns,

I now aim to show that—at least on the plausible conception of moral inquiry as a collective enterprise—accepting the epistemic significance of agreements is in fact crucial. This culminates in a reversal of the initial charge against moral conciliationism. I argue that epistemological moral scepticism can *only* be avoided if we assign moral agreements substantial epistemic significance. If so, then rejecting moral conciliationism may very well lead to epistemological moral scepticism.

Let me begin on an anecdotal note. Contemporary academic moral theorising is quite obviously a collective enterprise. All standard methods of research involve more than just an individual ethicist. We read about other’s moral views, we respond to these published views, we incorporate aspects of their views into ours. We present our views at conferences, hoping for comments on their plausibility and ultimately aim to publish our research to an interested audience. Presumably, all of this is done not merely because it is our job or because we expect some form of eternal glory. Really, we are interested in what other people think about our position, hoping to learn from them, etc. Our practices of moral theorising are inseparable from these social aspects.

Once we turn to morality ‘on the ground’, it becomes even more obvious why it would be a mistake to think of morality without this collective aspect. The moral domain is vast, moral questions are difficult and almost everyone participates in moral inquiry. Similar to the academic circles, we care about other’s moral views, how they perceive our moral convictions, etc. The vastness and the complexity of morality also mean that we quite often rely on others in making moral decisions. In short, both moral practice ‘on the ground’ as well as academic moral theorising are inextricably linked to certain social practices.

Of course, none of this conclusively shows that morality is necessarily a collective enterprise. We could, after all, be mistaken in our practice. If you think that we can figure out morality by ourselves, from our armchair as it were, these considerations will probably not convince you of the opposite. However, I am content with providing these considerations as a reason to think that it is at least extremely plausible that morality is closely linked to social practices. And insofar as we think that these practices have epistemic value, we should be sceptical about the claim that agreements (and indeed, many other forms of testimony) lack epistemic significance. After all, much of our moral practices either depend on or aim towards such moral agreement. If anti-conciliationism binds you to the armchair, all the better for my project here.

You might be tempted to claim that such agreements are of epistemic significance only if they provide us with new reasons for a particular view—it is not the pure fact of agreement, but rather the reasons themselves that do the epistemic work here. Agreements that provide us with new reasons are significant because of the underlying (new) reasons. It then comes down to the question

but rather due to *unusability* concerns. According to these, moral testimony ought to be avoided because actions based on moral testimony are *morally* (cf. e.g. Nickel (2001) or Hills (2009)) or *psychologically* (cf. e.g. Howell (2014), Fletcher (2016) or Callahan (2018)), but not epistemically, sub-optimal. In chapter 2, I argue that such unusability concerns ultimately fail.

whether the pure fact of moral agreement (i.e. those agreements that do not provide new reasons) can have epistemic significance. The task is therefore to argue that the pure fact of agreement can carry epistemic significance.

In the following, I assume a weak form of foundationalism for morality: some moral knowledge is basic, i.e. such that it can provide you with knowledge prior to knowing the reliability of the source, the prime example being moral intuitions (cf. Cohen (2002), 310). Since agreement on such intuitions provides a mere alignment of intuitions, there are no reasons over and above the agreement themselves to be gained.

Apart from contradicting our moral practices, assuming that agreements on moral intuitions are epistemically insignificant commits us to believing that our own moral intuitions are in some sense special, more likely to be right than those of others. For suppose you were asked to justify a particular moral intuition of yours. If such intuitions provide basic knowledge, then simply providing your intuition is, as a first step, sufficient. However, once you take into account that others may report conflicting intuitions you either have to give up on your intuition (because a justified challenge came up) or claim that you are nonetheless justified in the belief your intuitions supports. If you cannot make this latter move by drawing on agreement, the only alternative is to claim that your intuitions take some form of precedence over those from others. Seriously taking the relevance of agreements into account allows us to assign the same relevance to other's intuitions as to our own and simultaneously remain entitled to hold some of the corresponding intuitions (namely those with sufficient agreement). Really, then, we should have trust in our (*plural*) intuitions rather than our (*singular*) intuitions. Our trust in our individual intuition may only be justified derivatively.

I have argued that taking moral agreements to be epistemically insignificant is both in contradiction with our moral practices and likely to lead to counterintuitive consequences in the case of basic moral knowledge such as moral intuitions. On the understanding of morality as a collective enterprise that I have sketched taking moral agreements to be epistemically insignificant is implausible.

Taking this conceptualisation of morality as a collective enterprise seriously may not just render moral conciliationism plausible, but indeed *necessary* to avoid epistemological moral scepticism. For given the importance of the collective level, rejecting moral conciliationism leads us dangerously close to epistemological moral scepticism. Given symmetry, the anti-conciliationist must reject *both* the epistemic significance of moral agreement and moral disagreement. Accordingly, they cannot draw on the collective level to justify their moral beliefs (or at least, they cannot do so for *basic* beliefs as I explore them here—they can, of course, still draw on the *reasons* provided by other agents). Without drawing on collectively justified intuitions, however, we have no reason to think that *our* intuitions are the reliable ones. Without drawing on (dis)agreements, we are on our own even when it comes to a moral reality that is not of our own. Thus, we must remain sceptical if we attribute no epistemic significance to the testimony of others, including their agreements and, by symmetry, disagreements.

Hence, insofar as we think of morality as a collective enterprise, denying the epistemic significance of agreements on moral intuitions entails epistemological moral scepticism.¹⁴

Conciliationism in general and moral conciliationism in particular were initially introduced as positions on what to do in the face of disagreement. The same goes for its competitors in the anti-conciliationist camp. What the discussion in this chapter has shown is that this restriction in scope is too narrow. The positions in the debates around disagreement are plausibly not just understood as being concerned with disagreement, but with the relative value of other's moral testimony in general—including disagreement and agreement. It is this shift in focus that allows the defence of moral conciliationism I put forward.

Moral anti-conciliationism undercuts the justification for many of our moral beliefs by removing moral agreement as a justifying factor. Rejecting moral conciliationism seemed plausible at first because of how extreme moral conciliationism appeared. However, in accepting that moral beliefs must be justified entirely without reference to both agreement and disagreement, the anti-conciliationist goes even further. In particular, they lose moral intuitions as a basic kind of moral knowledge—a result that renders moral conciliationism a rather innocent position by contrast. If avoiding epistemological moral scepticism is desirable, adopting rather than rejecting moral conciliationism is the way to go.

1.4 Conclusion

In this chapter, I have argued that the charge of epistemological moral scepticism against moral conciliationism is unsuccessful. For this, I first established a symmetry claim of the epistemic significance of agreements and disagreements. Drawing on this symmetry claim, I argued that moral conciliationism offers a plausible way out of the suspension of judgment on our moral beliefs—agreements. So, moral conciliationism does not entail epistemological moral scepticism.

In the final section of this chapter, I further suggested that moral conciliationism may indeed be *necessary* to avoid epistemological moral scepticism. I suggested that moral inquiry is best understood as a collective enterprise. Insofar as you agree with this conception, we should assign moral agreements epistemic significance. Symmetry then commits us to assigning equal epistemic significance to moral disagreements.

¹⁴An anonymous referee has noted the similarity of my argument here to the argument Elga (2007) presents in defence of conciliationism in general. This argument has been subject to disagreement holding that it is okay (or even required) to form a fundamental trust into one's own mental states that is not appropriate for mental states of others (cf. e.g. Enoch (2010) and Wedgwood (2010)). If this criticism is successful, the arguments in this chapter show that it also has a further, perhaps undesirable, consequence. Used as a defence against disagreements, it gives the critic what they want. However, seeing that we similarly would have to reject the relevance of agreements with others on basic beliefs, it would also entail that any further agreement would not affect our level of confidence in the belief at hand (for our trust in our own mental states is much more fundamental).

Since there are two sets of premises doing the crucial work in my arguments, there are roughly two ways to resist my argument. First, one might reject the symmetry claim and argue in particular that disagreements lack epistemic significance while agreements do not. Second, one might accept the symmetry claim but argue that both moral disagreements and moral agreements are void of epistemic significance. I hope to have made a convincing case that neither of these options is particularly attractive.

Chapter 2

The Limits of Pessimism About Moral Testimony

Should we act on the basis of moral knowledge acquired purely through moral testimony? Many are pessimistic about this. Here, I argue that influential moral and psychological sources of such pessimism fail to support proper pessimism about moral testimony. Pessimism about moral testimony on moral grounds holds that testimony-based actions cannot be of moral worth. I argue that while there is some value in acting with moral worth, this value is negligible when compared to acting rightly. Pessimism about moral testimony on psychological grounds attests that testimony-based moral beliefs lack motivational and affective states accompanying ‘regular’ moral beliefs. I argue that the motivational aspect can be maintained if deference occurs out of respect for one’s own limitations. The only problem that testimony-based moral beliefs retain is their lack of corresponding affective states. Yet, this does not entail the illegitimacy of deference in moral matters either—reliable ‘cold’ moral beliefs are often better than less reliable ‘regular’ moral beliefs. Thus, I propose that a moderate optimism about moral testimony is appropriate: even if there is something suboptimal in reliance on moral testimony, this rarely provides a sufficient reason to reject reliance on moral testimony. Crucially, what feels suboptimal about deferring is more often than not something suboptimal about the environment an agent finds themselves in, but this non-ideal environment is exactly why the agent should defer.

2.1 Introduction

Suppose you visit a friend in a town with which’s local politics you are unfamiliar. One morning, you find your friend gone off to work. They left a note saying “We’re having this really important protest against the construction of an office block at central station at 4pm.” Arguably, you can legitimately infer from the note that there is a protest at central station at 4pm. However, it seems somehow illegitimate for you to *join* the protest merely based on their testimony. Intuitively, the legitimacy of reliance on non-moral and moral testimony differs.

We can substantiate this intuition by considering the interactions you might have if you decide to join the protest. Arriving at the protest, a TV crew asks you why you oppose the construction—you give your friend’s testimony as the reason. In a report about the protest, the moderator notes with some amusement that some of the protesters did not even understand *why* they were protesting.

Later that day, you spot an acquaintance of your friend. They come over and tell you how grateful they are that you joined the protest. After all, the matter is of great importance to them. When they ask why you joined the protest, you give your friend’s testimony as the reason. Upon learning this, they turn around, mumbling “oh, then you can’t *really* care” and walk away.

If you have these or related worries about reliance on moral testimony, you are in good company. Many philosophers agree that something about reliance on testimony in moral matters seems suspect. They diverge, however, on whether this apparent illegitimacy has any substance to it or whether our initial intuition can be explained away. Optimists about moral testimony maintain that reliance on moral testimony is, despite its illegitimate appearance, perfectly fine. In contrast, pessimists about moral testimony hold that there is in fact something illegitimate in relying on mere testimony when it comes to moral matters. Of course, there are varieties in the strength of such pessimism. On the one end, we have pessimists such as Nickel who hold that “it is not possible” for an action based on moral testimony “to be morally good” (Nickel (2001), 65). Most pessimists, however, are more moderate, holding that while an action based on moral testimony can be morally good, it usually is not—due to a variety of features of moral testimony (cf. e.g. Hills (2009), Howell (2014), Fletcher (2016), and Callahan (2018), all of whom I focus on in this chapter).

In this chapter, I focus on these more moderate positions—the thought that there is something suboptimal in reliance on moral testimony such that we usually should not rely on moral testimony. I argue that the motivations these pessimists have brought forward fail to support proper pessimism about moral testimony. In particular, I argue that *even if* there is something suboptimal about moral testimony, this rarely provides a sufficient reason to reject reliance on moral testimony. We thus end up with an optimistic position about moral testimony—albeit a moderate one. This strategy of proposing a moderate optimism about moral testimony by acknowledging some ‘problems’ with moral testimony but arguing that they do not entail that moral deference thus be-

comes illegitimate is due to Lord (2018). Lord identifies a lack of acquaintance with ethical facts as the most promising motivation for pessimism about moral testimony, and argues that this does nonetheless not entail that deference is illegitimate. I apply this strategy to the more commonplace worries pessimists have raised against reliance on moral testimony, thus broadening its scope.

First, I discuss moral reasons for pessimism about moral testimony. Pessimism about moral testimony on moral grounds holds that testimony-based actions cannot have *moral worth* because the agent in question lacks moral understanding. I argue that while there is some value in acting with moral worth, this value is negligible when compared to acting rightly.

Second, I discuss psychological reasons for pessimism about moral testimony. These include a lack of *motivational* and *affective* states in testimony-based moral beliefs which usually accompany moral beliefs. I argue that the motivational aspect can be maintained if deference occurs out of respect for one's own limitations. The only problem that testimony-based moral beliefs retain is their lack of corresponding affective states. Yet, this does not entail the illegitimacy of deference in moral matters either—reliable ‘cold’ moral beliefs are often better than less reliable regular moral beliefs.

In proposing a moderate optimism about moral testimony I hope to offer a place of reconciliation for pessimists and optimists. Morality is vast and none of us can specialise in all its aspects. Hence, we may often require guidance from others in morality. If so, then it is helpful if reliance on testimony is not as problematic as previously held all the while keeping its potential drawbacks in mind. Pessimism about moral testimony often points to the non-ideal nature of the deferring agent, and should—as I argue later—just as often point to the non-ideal environment an agent finds themselves in. But, I suggest, such non-ideal circumstances are exactly those where deference becomes viable and indeed mandated.

In focusing on moral and psychological sources for pessimism about moral testimony, I assume that moral testimony is sufficient to provide knowledge. Thus, I presuppose that worries of an *epistemic* nature do not succeed in rendering reliance on moral testimony illegitimate. Such worries include problems in identifying reliable testifiers in the moral domain (cf. McGrath (2009)) or the high stakes in moral matters that might render reliance on moral testimony insufficient for knowledge (cf. Hopkins (2007)).¹ This is in line with most of the literature on optimism and pessimism about moral testimony.

¹Accordingly, my discussion falls into the class of *unusability* concerns. These hold, in contrast to *unavailability* concerns, that moral testimony makes moral knowledge available but renders it unusable (cf. Hopkins (2007)). In distinguishing between moral and psychological sources of such unavailability, I show that a simple identification of unusability accounts with moral sources and unavailability accounts with epistemic sources is inadequate. Unusability accounts encompass more than moral sources (namely psychological sources), unavailability accounts encompass also metaphysical sources, most prominently discussed in McGrath (2009), but also, if only indirectly, in Cath (2019).

2.2 Moral Understanding

Recall the first problem you could encounter upon joining the protest purely on the basis of your friend's moral testimony. Even if it is the right thing to protest, you fail to understand *why* it is. This is due to the nature of testimony. Mere testimony can only be a non-explanatory reason which does not (by itself) allow for understanding: knowing that some action is morally right on the basis of moral testimony does not by itself allow you to understand why it is. For such understanding, explanatory reasons—those reasons that not only justify your belief, but *explain* the believed proposition—are required. Those who maintain that the problem with moral testimony is of a moral nature attest that this lack of understanding explains why testimony-based moral knowledge is unusable.²

This explanation is, however, not self-standing. The lack of understanding in testimony-based beliefs is not unique to the moral domain. Your non-moral belief that there is a protest is also based on your friend's testimony. As with the moral belief, their testimony does not enable you to (fully) understand why the protest takes place. Nonetheless, deferring on this non-moral matter seems legitimate. Given that the lack of understanding is due to testimony's failure to provide explanatory reasons, the question is whether there is a special connection between explanatory reasons and moral beliefs that does not exist for non-moral beliefs.

Thus, the pessimist about moral testimony still needs to show why understanding is particularly valuable in the moral domain. Before moving on to such proposals, let me briefly give an account of understanding as the common ground for the following discussion. Hills (2009), a pessimist about moral testimony herself, defines understanding in terms of abilities:

If you understand why p (and q is why p), then in the right sort of circumstances, you can successfully:

- (i) follow an explanation of why p given by someone else;
- (ii) explain why p in your own words;
- (iii) draw the conclusion that p (or that probably p) from the information that q ;
- (iv) draw the conclusion that p' (or that probably p') from the information that q' (where p' and q' are similar to but not identical to p and q);
- (v) given the information that p , give the right explanation, q ;
- (vi) given the information that p' , give the right explanation, q' .

To understand why p , you have to have the abilities i–vi to at least some extent. (Hills (2009), 102–103)

²Jones (1999) mentions a similar worry but presents it as a reason to think that moral testimony does not suffice to provide moral knowledge. On her interpretation, the pessimist might argue that the primary source for moral knowledge are moral reasons which are not provided by pure testimony. Hence, moral testimony would fail to transmit moral knowledge (cf. Jones (1999), 57–59).

Whether understanding is best spelled out in terms of abilities is subject to debate. My discussion in the following is compatible with alternative notions of understanding (cf. Ross (2020) for an overview of some such notions).

Hills holds that understanding is factive, i.e. you cannot understand p if p is false (cf. Hills (2009), 99). However, while she holds that knowledge is possible without understanding, she is not explicit about the reverse relation. She argues that moral understanding is not merely knowledge plus some other pieces of propositional knowledge since it involves abilities that cannot be reduced to propositional knowledge (cf. Hills (2009), 103). However, it seems as if someone who understands why p will usually also know that p on Hills’s account. My point here is not that understanding is in some sense reducible to knowledge (Hills’s notion of understanding in terms of abilities straightforwardly excludes this option). It is merely that—at least typically—when you understand why p , you know that p .³ After all, by (iii), you are able to conclude p from q which, in order for you to understand why p , you usually know (there may be non-standard cases in which this does not hold, e.g. those where you have merely implicit knowledge that p).

When considering whether the apparent illegitimacy of relying on moral testimony is connected to its failure of providing us with moral understanding, it will be helpful to keep this in mind. If understanding typically comes with knowledge, then there is at least one way in which understanding is usually preferable to knowledge—it simply gives the agent ‘more’.

2.2.1 Moral Worth

Hills presents three motivations for valuing moral understanding, some of which are already gestured at in Nickel (2001) (cf. Hills (2009), 106–119., but also Hills (2013), 555):

- (a) Moral understanding enables us to *reliably* do the right thing.
- (b) Moral understanding enables us to *justify* our actions to others.
- (c) Moral understanding enables us to perform *morally worthy* actions.

All of these motivations concern not just beliefs, but actions. This raises the question whether the illegitimacy of deferring to moral testimony results from *believing* on the basis of testimony or *acting* on the basis of testimony-based beliefs. Howell (2014) notes that both doxastic and active moral deference (as he labels these options respectively) seem illegitimate. In the following, I assume that this observation captures our intuitions about moral deference. Still, it is worth explicating that some sources of pessimism will be more closely aligned with doxastic deference and others with active deference. Given that

³Ross (2020) convincingly shows that understanding does not always entail knowledge as understanding cannot be ‘gettiered’ in the same way that knowledge can. Still, I take it that the connection between understanding and knowledge is as outlined above—typically, if you understand why p , you also know that p .

all of Hills's motivations involve actions, it seems reasonable to understand her account as being concerned with more than just doxastic deference—namely, active deference.⁴

Before moving into the individual discussion of (a) to (c), let me note one problem all of these approaches face but which I set aside for the purposes of this chapter. The pessimist about moral testimony needs to show not just that moral understanding is *preferable* to testimony-based moral knowledge, but that testimony-based moral knowledge itself is best to be avoided. For this argument to succeed, reliance on moral testimony would have to *prohibit* us from gaining moral understanding. But moral testimony seems compatible with and no obstacle to moral understanding. Indeed, moral testimony may sometimes *promote* the development of moral understanding (cf. Sliwa (2012), 192 and Callahan (2018), 443). While I do find this line of argument convincing, I think more can be said about the motivations Hills presents.

The idea behind (a) is that while moral knowledge based on moral testimony might enable us to do the right thing in a particular situation, it fails to enable us to do the right thing in related situations. Reliance on mere testimony comes with a failure to grasp the relevant moral reasons underlying the right action. Hence, testimony-based moral knowledge will usually be isolated. In contrast, by definition, moral understanding enables us to do the right thing also in related situations because we understand the moral reasons underlying our action and can draw inferences on the moral reasons in related situations (cf. (iv) in the definition of understanding above).

However, (a) fails to support an asymmetry between moral and non-moral testimony. Just as with moral testimony, deferring to others in non-moral matters also provides us with isolated knowledge and leaves us unable to draw conclusions about related subject matters. The isolated nature of testimony-based knowledge is merely an explication of the earlier insight that understanding typically comes with knowledge and is, in this sense, preferable to mere knowledge. However, this does not entail the unusability of testimony in the non-moral domain and it does not provide a reason to think that the moral domain should be special in this respect.

What about (b)? Let us suppose that justifying our actions to others is important when it comes to morality. After all, we do so on a regular basis and frown upon those who cannot justify their actions in this way.⁵ Furthermore, moral understanding arguably enables us to provide a justification for our actions—the reasons we take to support our belief (cf. Nickel (2001), 256). If we grasp the moral grounds for our beliefs and actions, we can, at least in

⁴This makes a difference only if you believe that assertion and practical reasoning do not have common epistemic standards. Brown (2012) provides a helpful discussion of why some apparently plausible commonality claims fail. Whether or not such a commonality claim holds, I focus on the doxastic side of things throughout my discussion.

⁵In their treatment of reasoning more generally, Sperber and Mercier draw on empirical evidence to suggest that the *primary* function of reasoning—including moral reasoning—is justifying our actions to others (cf. Sperber and Mercier (2017)). For a normative approach, take Scanlon's contractualism which has the justification of our actions to others as a central tenet (cf. Scanlon (1998), 147-188.).

principle, provide others with the demanded justification. There are, however, two problems with this proposal.

First, recall that the explanation the pessimist provides for the value of moral understanding should not entail that non-moral understanding is valuable in the same way. For if so, they lack an explanation for why reliance on non-moral testimony is legitimate. However, justifying our actions is also of relevance when it comes to the justification of our non-moral beliefs. What is more, it seems likely that the importance of the justification will increase relative to the importance of the action in question. Thus, the asymmetry would be between important and unimportant beliefs rather than moral and non-moral beliefs. Still, let us assume that these two distinctions correlate such that most moral beliefs will be important. Even if this explanation can account for the asymmetry, is the explanation convincing?

This is where the second problem with the proposal comes in. For this motivation to explain the value of moral understanding, we need to assume that we can only legitimately justify our actions by appeal to *explanatory* evidence (as opposed to non-explanatory evidence) in moral matters. However, assuming that non-explanatory evidence such as mere moral testimony is illegitimate begs the question (cf. Wiland (2021), 67 for a similar point). After all, we are trying to *find out* whether moral testimony is legitimate to draw on for our beliefs. Simply assuming that moral testimony cannot be a good justification for actions does not help. Recall the introductory example of a protest which you join solely on the basis of someone else's testimony. Even if you think that their moral testimony is not the proper justification in this case, this is a non-trivial finding and requires further explanation. The pessimist about moral testimony has yet to show what it is about the moral domain that renders explanatory reasons such an important role that reliance on mere moral testimony is illegitimate.

Let us therefore turn to the final motivation Hills provides, (c). This suggestion is in line with proposals about the moral worth of actions in a broadly Kantian tradition. According to so-called right-reasons views, an action is morally worthy if it is the action that ought to be performed morally and if it is performed for the normative reasons which it ought to be performed for (cf. Markovits (2010), 205).⁶ On such accounts, actions performed merely on the basis of moral testimony never have moral worth.⁷

⁶Markovits distinguishes two readings of the Kantian idea of acting for the right reasons. First, the motive of duty thesis according to which actions must be performed because they are right. Second, the coincident reasons thesis according to which actions must be performed for the reasons why they are right. Markovits rejects the motive of duty thesis in favour of the coincident reasons thesis (cf. Markovits (2010), 202–207).

⁷This reveals why pessimists about moral testimony on moral grounds are best understood as being concerned with active deference. After all, *doing* the right thing for the right reasons is inextricably linked to action. There are two ways the pessimist of this kind might resist this classification. First, they might hold that the account of actions simply transfers to beliefs. On such a view, beliefs would have moral worth only if they are believed for the right reasons. Second, the pessimist may hold that the moral domain requires an intimate link of belief and action such that every worry about illegitimate action entails similar problems for the underlying beliefs.

This final motivation does not fall prey to the shortcomings of (a) and (b). First, moral worth seems to provide a good reason why explanatory reasons play such a central role in the moral domain.⁸ Second, it also explains the asymmetry between moral and non-moral testimony—after all, there seems to be no such thing as ‘prudential worth’ and even if there were, it seems much less important than moral worth (cf. Fletcher (2016), 55). If you find your way through town merely by relying on your maps app, it does not seem that you thereby forego a relevant achievement that you would have gained had you developed an understanding of the town instead.

To establish proper pessimism about moral testimony, then, the relative importance of performing morally worthy actions has to be of significant value. Otherwise, it fails as an explanation for the illegitimacy of relying on moral testimony as it does in the case of prudential deference. If the value of moral worth is minor, it cannot do all the work the pessimist requires of it. This does not entail that the pessimist about moral testimony needs to hold that the value of moral worth *always* outweighs other considerations. Rather, the value of moral worth ought to provide a strong—but defeasible—reason to opt for contemplating on our own instead of deferring. In this framework, we can distinguish three factors that determine whether or not you ought to defer (cf. Hills (2009), 124).⁹

- (1) The likelihood of you acquiring moral understanding if you try on your own.
- (2) The likelihood of you acquiring moral knowledge if you defer.
- (3) The relative importance of performing morally worthy actions.

Given this setup, there seems to be an easy way to show that we should, other things equal, opt for figuring things out on our own rather than deferring. Suppose you had the option to learn about what to do in a particular situation either by contemplating the reasons for and against or by invoking the moral testimony of a reliable testifier. Suppose further that you know that the likelihood of getting things right is the same for either method. What should you do? Clearly, it seems, you should opt for contemplating on your own:

Forming a moral judgment by deference and then acting on it is much less
of a moral achievement than forming the true judgment without deference,

⁸Callahan argues that while moral understanding may be necessary for moral worth, it is not sufficient (cf. Callahan (2018), 445). You might well understand (in Hills’s sense) a moral matter without being motivated to follow through on your understanding. Callahan’s solution is to expand the notion of moral understanding such that it includes proper motivations. I discuss this approach in section 2.3.2.

⁹Of course, there are other factors that may influence whether or not you should, in a particular situation, defer. Most notably, this includes the cognitive costs of both processes—arguably, relying on moral testimony is usually easier than contemplating on your own. In situations with limited capacity, this may provide a practical reason to defer even if it turned out that moral understanding really is very valuable. I come back to such limitations in the final section of this chapter. Another important factor are the stakes involved in a particular decision. Arguably, the more important some matter is, the more important it is to get it right.

because it doesn't constitute the appropriate response to the morally relevant features of the case. (Enoch (2014), 256)

However, this does not yet show that the relative importance of performing morally worthy actions is significant. Deciding to figure things out on your own can be justified without ever drawing on the value of moral worth. As outlined above, moral understanding typically comes with moral knowledge. Given that moral understanding therefore usually gives you 'more' than moral knowledge, this is reason enough to prefer moral understanding over moral knowledge if the likelihood in (1) and (2) is equal. Moral worth itself does not yet play an explanatory role. All that follows so far is that the value of performing morally worthy actions is not negative—but that is not enough to establish proper pessimism about moral testimony.

To demonstrate that the relative value of performing morally worthy actions is significant, we need to show that even if testimony-based moral knowledge makes it more likely for one to do the right thing, one should *nonetheless* opt for contemplating rather than deferring because of the value of performing morally worthy actions. Let us consider such a potential case adapted from one by Nickel (2001):

Suppose my friend has asked me to loan them some money for rent, but I know that this may further support their irresponsible handling of money. I cannot decide what to do, and I ask my mother what she thinks. According to her, I ought to tell my friend that I will loan them some money only if they take a course on the responsible handling of money (which they do not want to do). In such a case, is it permissible for me to rely on my mother's word as my justification? What if my friend wants to know, as they have been evicted from their apartment, why I decided as I did? It seems doubtful to me that they would be satisfied by the reply, "That's what my mother told me I should do. She is much more experienced than I am at making decisions about this kind of thing, and she knows what to do." (adapted from a case by Nickel (2001), 256)

As we can see in this passage, Nickel himself thinks there is something illegitimate about deferring in cases like these. Still, I think there is a convincing case to be made that deference is in fact an acceptable, if not the required choice.

Suppose I have good reasons for thinking that relying on my mother's moral testimony is more likely to lead me to the correct belief and correspondingly to the correct action than contemplating on my own. If so, then my options are to either contemplate and hope that it will lead to the correct outcome thus allowing me to perform a morally worthy action or to defer, being fairly confident that my action will be the right one but also knowing beforehand that my action will lack moral worth. If this description is correct, then my friend should be very much on board with my decision to defer. After all, why should the potential moral worth of *my* action matter more than *their* well-being? Far from being the morally virtuous option, relying solely on my own contemplation is rather selfish. The potential moral worth of my action goes no way in improving my friend's situation. In this case, we can assume that the

likelihood of (2) surpasses that of (1) just slightly. If the value of performing morally worthy actions, (3), were significant, it should follow that we can opt for figuring the matter out on our own. But this seems wrong: taking a moral risk at the cost of someone else merely to our own benefit is not acceptable.

Considering the above case but with the alternative decision helps to illustrate this. Suppose I decide to ignore my mother’s advice. I loan my friend the money, ultimately resulting in them not learning how to handle money responsibly for many years. At a later point, we reflect on the situation. When asked to justify why I did not simply defer to my mother, I reply: “I figured that deferring would have made it impossible for my action to count as an achievement—deferring would have meant that my action could not have had moral worth.” Should my friend really be satisfied with this response? The response is clearly negative.¹⁰ The potential moral worth of an action does not in itself outweigh the action’s potential to be the right action. The value of moral worth is not significant enough to explain the appearance of reliance on moral testimony as illegitimate. The decision to rely on someone’s moral testimony is thus best understood as one possible outcome of one’s deliberation on how to act in a given situation—the first step in making a moral choice is to decide on the best process; this may be own deliberation, but it may just as well be reliance on moral testimony.

2.2.2 Moral Worth Fetishism

Both the motivation for and my response to (c) come out even more poignantly in terms of a debate related to the debate around moral worth, but spelled out in terms of the proper moral motivation—an explicit concern with acting rightly on the one hand and being motivated by the right-making features themselves.¹¹ Consider Smith’s characterisation of what moral motivation ought to be like:

Good people care non-derivatively about honesty, the weal and woe of their children and friends, the well-being of their fellows, people getting what they deserve, justice, equality, and the like, not just one thing: doing what they believe to be right, where this is read *de dicto* and not *de re*. Indeed, common-sense tells us that being so motivated is a fetish or moral vice, not the one and only moral virtue. (Smith (1994), 75).¹²

¹⁰In chapter 6, I discuss the question of who is to be blamed when you rely on someone else’s moral testimony and end up getting things wrong because of it.

¹¹Weatherson also notes the similarity between these two debates—his discussion is about partial normative deference where the resulting action fails to be motivated in the right way (cf. Weatherson (2014), 152).

¹²As Johnson King (2020b) points out, motivation by moral rightness *de re* and motivation by moral rightness *de dicto* are neither exclusive nor exhaustive and the discussions of them as contraries has sometimes hindered progress in the debate. The so picked out motivations, she argues, are “not the only games in town” (cf. Johnson King (2020b), 261). For the purposes of engaging with the present discussion, I follow the literature’s framing and consider only these two options: the concern with acting rightly on the one hand and motivations by right-making features on the other. However, in chapter 3, I explore some of the problematic consequences of this framing as well as ways out of it.

Smith (1994) problematises an explicit (and exclusive) concern with acting rightly, arguing that it alone does not require grasping the moral reasons for one's action—and pure moral testimony seems to be the prime example of just that.¹³ Following Smith's characterisation, such motivations not only lack moral worth but constitute a moral vice. This is the charge of moral fetishism.

Relying on moral testimony, or so the argument goes, is a kind of moral fetishism since the agent relying on moral testimony fails to care adequately about the moral grounds of their action—its right-making features. However, even if one accepts that failing to care in this way constitutes some kind of moral fetishism, the real question is whether this moral fetishism is in some sense *objectionable*.¹⁴

Given my response to Hills's argument against moral testimony, you can probably guess where my response is going. Again, while it may be better to do be motivated by the right-making features directly rather than by a concern with acting rightly, the question is whether this difference in value is such that a mere concern with acting rightly is worse than attempting to do the right thing for the right reasons even if the former makes it more likely that you are in fact doing the right thing. And here, the answer is negative. If you care so much about the proper motivation that you are willing to risk failing to do the right thing in the first place, *then* you seem to be exhibiting an objectionable kind of moral *worth* fetishism. It seems rather dubious to think that the potential moral worth of an action (and, relatedly, the potential praiseworthiness of the agent) should ever justify risking to do the wrong thing. Especially considered from the first-person perspective, this kind of reasoning will almost always be selfish or at least reckless (cf. Sepielli (2016), 2961 for the latter point).

Of course, an ideal agent with perfect knowledge of both the empirical and normative landscape will never face such a situation. But this does not change the fact that most of us are, whether we like it or not, non-ideal. Suppose such a non-ideal agent finds themselves in a situation in which they recognise their own limitations in terms of properly responding to the moral features of a situation. I think Smith is right in thinking that there is something regrettable about our agent—it would have been better for them to have the capacity of responding correctly to the relevant features. But this does certainly not entail that we should blame them for recognising their limitations and choosing to

¹³An anonymous referee rightfully pointed out that not all cases of moral testimony will be as pure as those that pessimists about moral testimony are usually interested in. E.g. McShane and Wiland argue that in typical cases of moral testimony, for the most part, agents deliberate on their own coming up with reasons in favour or against the action. Their problem is not in identifying these reasons, but rather in weighing them competently. If moral testimony comes in only to resolve the question of weighing, it seems that the agent does grasp the moral reasons for their actions (cf. McShane (2018), 254–255 and Wiland (2021), 55–59). This seems correct to me—but many pessimists might be inclined to classify this as mere 'higher-order' moral testimony and worry that it does not do away with the usual problems of moral testimony—in particular, that it does not enable its recipients to do the right thing in related situations.

¹⁴Though e.g. Lillehammer (1997) argues that Smith fails to establish the charge of moral fetishism in the first place.

defer instead. A mere concern with acting rightly may in some sense be only the second-best option. But that does not entail that it is a *bad* option.¹⁵

In short, I agree with Smith that motivations by the right-making features may in some sense be preferable to a mere concern with acting rightly. However, this holds only if these two options have the same likelihood of enabling you to do the right thing. Once it is more likely that you would do the right thing if you rely on moral testimony, it would be objectionable to continue trying to track the right-making features instead of choosing the ‘safe’ route.¹⁶

So far, I focused on the potential negative consequences of risk-taking in moral matters where a ‘safer’ option is available. The risk of failing to do the right thing in the first place was most prominent. However, one may hold that there is at least something valuable in failing to do the right thing—as long as this failing stems from one’s own deliberation and not from the acceptance of someone else’s testimony. From such moral mistakes of our own, it might be argued, we learn a lot about morality while the same does not hold for moral mistakes resulting from someone else’s testimony. Did I neglect the value of moral mistakes in focusing on doing the right thing?

If it is indeed the case that we learn more about the correct moral choice by failing to do the right thing following our own deliberation than by failing to do the right thing after deferring, then it seems that there is a case to be made for running a risk in some cases. For an example of this, consider children whom we sometimes could aid in moral decisions but decide not to with the hope of them coming to learn the relevant moral facts on their own. However, this example also shows that the value we see in such moral mistakes need not stem from the potential of moral worth—if the stakes are high, we will not run the risk of letting them make a moral mistake. Risking not to do the right thing is thus ultimately not justified by the potential for moral worth but by the possibility to increase the chances of doing the right thing in the future. Thus, it is not the significance of moral worth that justifies running a moral risk here and there, but the likelihood of doing the right thing—where this is not understood as an one-off issue but as a more global issue.

¹⁵I am not taking a stance on whether or not agents motivated by a mere concern with acting rightly *can* act well. For now, my point is that even if they cannot, there may be good reasons to prefer such motivations over risk-taking behaviour. Cf. Hlobil (2020) for an argument that unproblematic cases of reliance on moral testimony coincide with cases in which an agent can act well when they are motivated by a concern for acting rightly (Hlobil (2020), 993–994).

¹⁶Sliwa (2012) even goes one step further. So far, I have argued that deferring is, under the circumstances just outlined, the right course of action because it increases the likelihood of performing a morally right action. Sliwa argues that choosing to defer in such scenarios is essentially doing the right thing (deferring) for the right reasons (resolving uncertainty about what to do), thus rendering deference to have moral worth (cf. Sliwa (2012), 193 and also Wiland (2021), 62–65 for a similar point). While I am sympathetic to this line of thought, it is important to note that it does *not* render the resulting action based on the moral testimony as morally worthy. It is only the act of deferring itself that can be of moral worth.

2.2.3 Moderate Optimism

Throughout the previous sections, I have characterised the pessimist as being so concerned with the downsides of moral testimony that they would rarely suggest forming a belief based on it. Admittedly, this characterisation may have been too strong. Maybe all the pessimist's position amounts to is this:

[E]ven when it is most reasonable overall to accept moral testimony, there is nevertheless *something* to be said for not doing so and for making one's own decision. (Crisp (2014), 134, author's emphasis)

However, if this is all the pessimist attitude amounts to, this *something* must be minor. Recall Hills's formula for whether or not you ought to defer. It consisted of the likelihood of acquiring understanding via own contemplation, the likelihood of acquiring knowledge via testimony, and the relative importance of performing morally worthy actions. If even the slightest increase in the likelihood of acquiring knowledge via testimony justifies deferring, then moral worth can hardly be said to have a significant value. If so, then the moral sources of pessimism about moral testimony do not support proper pessimism about moral testimony. Perhaps, then, there is a place for (some) reconciliation.

For if the optimist admits that the pessimist is correct in thinking that there is something to be said against moral testimony when compared to deliberating on one's own (namely, gaining moral understanding is preferable to moral knowledge) while the pessimist admits that this difference is relatively insignificant so that it is outweighed once the likelihood in getting it right via testimony increases, the difference between the two camps shrinks significantly. This way of reconciling the two camps can be understood in at least two ways. First, it may be understood as a way of re-framing the debate showing that self-declared optimists and self-declared pessimists are actually not as far away from each other as previously thought. Second, it can be understood as a new position for which both optimists and pessimists need to make some compromise. Ultimately, which of the two options is correct will depend on the strength of a particular optimistic or pessimistic position. For some, my suggestion might be nothing more than a reminder of the commonalities with the other camp and adopting it might not change much about their initial position—if so, then my proposal mostly offers a new framing of the debate focusing on commonalities instead of differences. For others, adopting my suggestion will require giving up on some of their prior commitments.

Either way, if the observations in this section are correct, then the main dispute comes down to the question *how often* the problematic feature of moral testimony has the effect that moral testimony should be rejected (cf. Callahan (2018), 440). I have provided reasons to think that moral worth itself will rarely provide a sufficient reason not to rely on moral testimony. In particular, it will only ever be relevant when it serves as a tiebreaker between contemplation and deference when these options have the same likelihood of leading you to the correct moral belief. Taking up this position, I submit, means adopting a moderate optimism about moral testimony: accepting that there are some

unfortunate drawbacks to reliance on moral testimony but that these rarely provide a reason against reliance on moral testimony.

So far, I have argued that deferring is the permissible or even required course of action if the likelihood of doing the right thing as a result of deferring is sufficiently high. If so, it may seem that the problem we have with moral testimony is, ultimately, not about the act of deferring itself but about something else. Some pessimists who accept this result turned to psychology-flavoured unusability accounts instead.

2.3 A Problem in the Agent?

If the source of the problem for moral testimony is not of moral nature, what else might it be? Recall the case of deference I introduced at the very beginning of this chapter. One of the negative responses to your joining the protest was by an acquaintance who thought that your reliance on moral testimony indicates a lack of care about the issue. Instead of a cognitive failure on your behalf, as evidenced in the moral understanding/worth explanation where the identified problem is the lack of a connection between the testimony-based moral belief and the moral reasons for the truth of the belief, we might instead look for *psychological* problems in your deference.

There are two related problems of this nature identified in the literature. The first is still closely related to the moral worth explanation I discussed in the previous section. Following this line of thought, reliance on moral testimony indicates a problematic character of the agent. The testimony-based moral belief fails to be properly integrated in the agent's character. After responding to this charge, I turn to the second psychological problem about moral testimony: reliance on moral testimony (usually) precludes the deferring agent from forming motivational and affective states.

2.3.1 Non-Integrated Moral Beliefs

Let us turn to the issue of the problematic character first. Relying on moral testimony, or so the idea goes, is not in itself objectionable. Rather, it indicates something problematic about the agent's *character*. Howell's explanation of this problem is nonetheless surprisingly similar to the moral worth explanation I considered above:

The belief of the agent and the actions it gives rise to will not necessarily redound to the agent's character because they don't stem from a virtue of the agent. Because the deference fails to bring with it feelings, intuitions and motivations, the resulting belief does not give rise to a virtue or reinforce existing virtues in the way the belief might if it were gained in other ways. (Howell (2014), 408)

The problem, as Howell identifies it here, is that testimony-based moral beliefs are not integrated into one's character. Hence, actions resulting from them

cannot reflect a moral virtue underlying them. This lack of an appropriate response unifies accounts explaining the asymmetry between non-moral and moral testimony via a psychological source. The problem, on this psychological explanation of the asymmetry, is not so much with the act of deferring but with the agent themselves. The lack of an appropriate response, Howell holds, indicates that there is something deficient about their moral character.

I think that this analysis is too one-dimensional—there may be other reasons than a deficient moral character that explain the deference. Before turning to such reasons in the final section of this chapter, let me argue that *even if* the proposed analysis *were* complete, it would not entail that deference in moral matters should often be prohibited.

Suppose an agent defers because they lack the ability (or think they do) to answer a particular moral question on their own. As Howell notes, this may indicate something suboptimal about the agent. However, it does not entail that there is something objectionable about their decision to defer. Granted, they fail to respond appropriately to the moral features of the situation they are presented with. However, to investigate whether there is something objectionable about the scenario, we should compare the agent in question not to a less flawed agent who responds on their own and appropriately to the moral features of the situation. Rather, the correct comparison class is a counterpart of our agent who shares the same moral ‘suboptimal’ features but chooses not to defer and instead tries to figure it out on their own. Even though this agent is aware of their limitations with respect to the moral issue at hand and is also aware of the option to resolve their uncertainty by deferring, they choose not to. They know that they might not only fail to act in the ways that the moral reasons support but may (and not improbably so) also fail to do the right thing in the first place (cf. Hicks (2019), 785).

It is for this reason that moral deference should not worry us. If Howell’s psychological explanation is right, then moral deference reveals the non-ideal nature of the deferring agent, that much is correct. However, *especially* if it reveals such a non-ideal nature, we should want the agent in question to defer. Compare this to a situation in a supermarket in which you are unable to find a product you desire. Since you’d really like to find the product, you consider asking the store assistant for help. However, just as you’re about to ask, it occurs to you that asking for help would indicate a failure on your behalf—namely, that you didn’t locate the product yourself. But of course (and only if we consider this to be a failure at all) it is not *asking for help* that is the failure. In fact, asking for help is your best means of resolving your failure to locate the product.

Howell seems to recognise this as he rightly points out that the deferring agent’s non-ideal moral nature does not *result* from deference but is *revealed* by it. This, however, should lead us to conclude that the deferring agent chooses the second-best option because, by assumption, it is unlikely that they will be successful in aiming for the best option (reliably finding the truth on their own). We might conceptualise this as a conditional obligation. If one is aware of one’s non-ideal moral nature, then one should opt to defer. Accordingly, when

faced with someone who defers on a moral matter for good reasons, nothing is going wrong. Given their options, they chose the right course of action—acknowledging that someone else is in a better position than ourselves to form a particular moral judgment (cf. Sliwa (2012), 179).¹⁷ Thus, even if the pessimist is successful in identifying the morally flawed character of the agent as the suboptimal feature in reliance on moral testimony, they do not thereby establish proper pessimism about moral testimony: *especially* in light of a morally flawed character, deference becomes a viable option.¹⁸

Perhaps the defender of the psychological explanation can admit as much about such agents. The problem, they might hold, arises from the fact that moral deference usually does not reflect such a self-critical or otherwise constrained agent. Rather, the real psychological problem is with a ‘lazy’ agent, who outsources the moral labour not because they recognise their own limitations but because they lack the drive to attain moral beliefs on their own (cf. Howell (2014), 403 for this characterisation and Wiland (2021), 59 for some critical discussion of this proposal). There are two problems with this proposal. First, it relies on the empirical assumption that moral deference usually occurs because of moral laziness instead of the recognition of one’s limitations (I give some reasons to doubt this in the next section). Second, even if this empirical assumption holds, it is far from obvious that deferring is such a bad option after all. If the psychological explanation is right, then the best option, becoming morally ‘active’, may currently be unavailable for our agent. Changing our moral and intellectual dispositions cannot be done overnight—it takes time. But if the best option is currently unavailable, then the second-best choice for our ‘lazy’ agent is to defer. Whether they defer or not, the option to become morally active *afterwards* remains an option.

In a nutshell, then, there is some truth to this version of the psychological explanation—it can explain our unease about agents deferring in moral matters. However, the explanation lacks the *normative* upshot the pessimist about moral testimony requires. Quite the contrary: for agents who lack either the ability or drive to gain a particular bit of moral knowledge, deferring can be better than trying on their own.

2.3.2 ‘Cold’ Moral Beliefs

So far, I have argued that even if deference in moral matters reveals a non-ideal moral character, there is nothing wrong with deferring for agents with a non-

¹⁷Assuming that the agent is limited in their ability to respond correctly to the first-order moral question it may seem likely that they will also be unable to tell who is a reliable testifier on the matter. If so, this raises an epistemic problem to moral testimony akin to the discussion in McGrath (2009). Roughly, the usual explanation for our ability to know who is reliable in responding to a question without being able to respond to the question properly ourselves draws on so-called independent checks. However, it seems that no such independent checks exist for the moral domain (cf. McGrath (2009), 334).

¹⁸Given Howell’s self-characterisation as suggesting an account focused on virtues, one might add that the virtue of humility, too, must feature in the account. In the situations just outlined, this virtue of humility would at least provide some reason to defer.

ideal character. There is, however, an alternative reading of the psychological explanation that does not require any claim about the moral character of an agent but focuses solely on the ‘cold’—and thus supposedly problematic—nature of testimony-based moral beliefs (cf. Callahan (2018), 446). These accounts hold that an essential part of a moral belief is its motivational and affective component—components testimony-based moral beliefs lack.

Unlike other beliefs, moral beliefs usually come with these desire-like aspects. The most extreme formulation of this idea is motivational internalism holding that there is a necessary connection between a moral belief and corresponding motivational (and presumably affective) states. In short, you cannot sincerely believe that performing an act is good without being motivated to perform the act. But we need not assume motivational internalism to get the idea of ‘cold’ moral beliefs off the ground. The observation that our moral beliefs are usually accompanied by these desire-like components is sufficient for our purposes here. These components are best accounted for by the explanatory reasons underlying our belief. We are motivated to perform an act because we sincerely believe that and understand why performing the act is good. If so, then testimony-based moral beliefs may often lack these motivational and affective components because they are based on non-explanatory reasons.¹⁹

For the purposes of this chapter, I grant the pessimist about moral testimony as much—even though one might also try to show that the motivational and affective components of moral beliefs can stem from non-explanatory sources as well. In any case, this line of thought has the potential to explain the asymmetry between moral and non-moral testimony: there is no analogue to the motivational and affective components in non-moral beliefs.

Fletcher (2016) first suggested this version of the psychological explanation, arguing that the discrepancy of testimony-based moral beliefs and affective or motivational states explains why reliance on moral testimony is illegitimate. Callahan (2018) expands on this solution and argues that reliance on moral testimony not only fails to bring about these desire-like components but also disincentivises further inquiry into the matter, thus rendering the acquisition of motivational and affective states in the future unlikely.

In responding to these charges against moral testimony, it is helpful to distinguish the affective states involved from the motivation. Let us begin with the motivational component. Arguably, being motivated to do the right thing is valuable. Therefore, a lack of motivation in testimony-based moral beliefs might undermine their value significantly. Consider, however, why people seek out moral testimony. For example, Sliwa suggests moral testimony as a way of resolving uncertainty, crucially while still being interested in how to act in light of the question (cf. Sliwa (2012), 193).

¹⁹If one is willing to accept motivational internalism and holds that moral testimony fails to transmit motivation, then one might end up with a metaphysically motivated pessimism about moral testimony. If moral testimony fails to transmit motivation, then testimony-based beliefs are, on the assumption of motivational internalism, not sincere moral beliefs. This line of thought would, however, fall into the scope of unavailability concerns—moral testimony fails to make sincere moral beliefs (and thus moral knowledge) possible.

Common to these reasons for seeking out moral testimony, then, is an interest in gaining guidance for our actions. If this interest precedes seeking out moral testimony, the agent seeking out moral testimony has at least some initial motivation to act in accordance with the testimony they receive. Compare this with a situation in which you want to visit a café in a city you are unfamiliar with. Not knowing the area, you ask a friend for a suggestion. Upon them suggesting *Combini Café*, you acquire a motivation to go there—albeit not for the reasons that actually explain why the café is a good choice. Similarly, testimony-based moral knowledge provides you with an indirect motivation to follow through on the testimony.

Now, as I noted in the discussion of moral worth, this indirect motivation cannot be a motivation by the right-making features (for lack of understanding). Nonetheless, it can be a motivation following a concern for acting rightly. The testimony-seeking agent finds themselves unable to figure out the right thing on their own. Still, they are motivated to do the right thing and therefore choose to seek out moral testimony. Upon gaining moral testimony, then, the agent may not acquire the moral belief's inherent motivational force. Yet, they will be motivated to act in accordance with the moral belief because they are motivated by a concern with acting rightly. Thus, if we think that the motivational component of moral beliefs is important solely because it produces agents who are motivated to do the right thing, we need not be sceptical about moral testimony.

Pressing this objection further one might of course argue that we care not only about being motivated simpliciter but about being *rightly* motivated—where this is understood as being motivated by the right-making features and not just a concern for acting rightly. If this is the objection, then my remarks about moral worth apply once more. For the objection to support pessimism about moral testimony, we would have to assume an objectionable form of moral worth fetishism. In conclusion, then, testimony-based moral beliefs do provide you with a motivational state to act in accordance with it, albeit via a non-standard route. If we instead care about the *right* motivation, our worry about moral testimony is resolved in the same way as the worry about moral worth was: caring more about being rightly motivated than being motivated to do the right things seems either selfish or reckless. Either way, the lack of inherent motivation in testimony-based moral beliefs does not render reliance on moral testimony illegitimate.

What about the affective states that usually accompany our moral beliefs? Surely these cannot be gained in a similar indirect manner as the motivational force can. I think this is correct. We might certainly point to individual cases in which your sympathy with the testifier leads you to develop an affective response even on the basis of pure moral testimony, but such cases are probably the exception.²⁰ In general, then, testimony-based moral beliefs will be lacking with respect to affective states, they will be 'cold'.

²⁰And even these exceptions seem somewhat odd. Suppose a friend of yours is mad at you because you acted in a way that they find objectionable. When inquiring about their reasons to find the act objectionable, they point to someone else's testimony. It seems rather natural to shrug off your friend's anger as inappropriate. Wiland (2021), 71–72, disagrees with this

Still, we should not dismiss moral testimony solely on the basis of this observation. Instead, we should analyse the reasons for the lack of affective state and consider the question of whether there are more desirable alternatives. If the answer to this latter question is negative, then there may be something *regrettable* about reliance on moral testimony, but it would not affect the *legitimacy* of relying on moral testimony in certain cases. Such a conclusion would further pave the way for the moderate optimism I suggest. We concede to the pessimist that there is something suboptimal about deferring in moral matters but argue that this rarely provides a sufficient reason against deferring.

Recall the argument for why deferring may be the correct response if an agent is incapable of getting to the moral truth themselves. In a similar vein, I take it that the choice to defer on a moral matter also indicates that the agent in question is, for some reason or other, incapable of forming the fitting affective states. Again, deferring will only be the second-best option. However, given that the best option is unavailable, deferring is exactly the right choice—though it comes with the regrettable drawback of lacking the affective states that usually accompany our moral beliefs. If so, then deferring is illegitimate only if there were some *additional* drawback in it. Indeed, deferring might at least ensure that the agent's belief/action is in some sense still based on the relevant emotional/motivational states, albeit those of another agent. If we care about the conative attitudes because these importantly inform moral decisions, it seems that deferring is the best way for our agent to ensure this connection.

Callahan's argument that deference disincentivises further inquiry into the matter at hand complicates this point, potentially offering the additional drawback deference brings with it. She argues that we are less likely to engage in inquiry if we already have a settled view on the matter:

My claim is that deference to testimony *dis-incentivizes*, and thus discourages the acquisition of understanding. For those with a settled view on a topic have, other things equal, less reason to engage in the kinds of reflective or experiential learning practices that might result in greater understanding of the matter. And the effect of deferring to moral testimony is to give hearers a settled moral view. (Callahan (2018), 454, author's emphasis).

If reliance on moral testimony closes off inquiry, then we will not only lack the fitting affective states *now* but are unlikely to form them in the future. This is, however, not an inherent problem of moral testimony—the real problem is with closing off inquiry at the wrong point. Consider an agent coming to a moral belief via their own contemplation. Given that they have now formed a settled view on the matter, they, too, have less reason than before to engage in the practices Callahan mentions. But this does obviously not entail that we should not contemplate on our own or form moral beliefs in general. Hence, Callahan is correct in thinking that it is possible to defer in a problematic manner, essentially closing off inquiry for the foreseeable future. But this does

evaluation—he thinks it is permissible to form affective responses based on e.g. the admiration of a testifier.

not entail that reliance on moral testimony in general is objectionable in this way. Rather, this insight should lead us to formulate rigorous constraints on moral deference—it is simply not as easy to defer as is sometimes indicated in the literature.²¹

Throughout most of the literature, moral deference is approached through the lens of supposedly lazy agents who choose to defer not for good reasons and in good ways, but for superficial and bad reasons. This, however, overlooks cases of moral deference going well. The moderate optimism I suggest is a place of reconciliation between optimists and pessimists about moral testimony. Reliance on moral testimony becomes a serious option only in non-ideal scenarios. But it is precisely these non-ideal circumstances that provide you with good reasons to defer.

2.4 Intertwining Dogmas

So far, I have argued that neither moral nor psychological sources of pessimism about moral testimony support proper pessimism about moral testimony. However, both the moral explanation as well as the psychological explanation continue to hold sway over most moral agents. The intuition that there is something wrong with deferring in moral matters is not easily refuted. Let me conclude by exploring one worrisome consequence these explanations entail.

According to the psychological explanation, deferring shows that there is something wrong about the agent, that they are, for some reason or other, incapable of coming to a moral belief on their own. Combine this with the assumption that figuring moral matters out on one's own is of special moral value and it becomes extremely costly (socially speaking) for an agent to defer in moral matters. After all, deferring indicates to others that the deferring moral agent is not living up to the ideal of a moral agent.

I take it that this result is undesirable. In obsessing over the moral worth of figuring it out on one's own, in obsessing over being rightly motivated, we seem to have forgotten that *all* moral agents are in some sense non-ideal. In light of this, we should take the value of recognising one's own moral limitations seriously in our theorising. Recognising that one is *not* properly positioned to respond to a moral matter is an important insight and certainly better than remaining ignorant about one's own shortcomings. Deferring in light of this recognition is definitely better in comparison to the alternative: stubbornly trying to figure it out on one's own. Deferring will probably not enable you to gain a new high score in the 'ideal moral agent' category, but you will still get a decent amount of points for recognising that you are not, in this respect, the ideal moral agent. By obsessing too much about what ideal moral agents ought to do, we effectively exclude non-ideal moral agents from accessing an important resource: relying on others.

²¹ Furthermore, it may often be advisable to close off inquiry—if only for practical reasons—whether or not after deferring to others or contemplating on one's own.

Throughout the literature there is another irritating assumption—namely that the failure of an agent to contemplate on a moral matter and attain a corresponding belief can only be due to a defect in their character or else due to some inherent flaw in the agent. Pressing the importance of non-idealised scenarios further, a crucial addition to this picture is that a failure to form appropriate moral beliefs on one’s own might also stem from social and other external factors. The obsession about moral worth inhibits discussions about the various ways in which an agent’s *environment* might be non-ideal.

In fact, this seems to be the most crucial oversight on behalf of most pessimists about moral testimony. The problems with moral deference they identify are almost always located *in the agent*. However, it seems very likely that the ‘problem’ is more often than not located outside the agent. In this vein, Jones and Kolers discuss how one’s social positions in society can give agents good reasons to sometimes rely on moral testimony (cf. Jones (1999), 77 and Kolers (2016), 103–112). Wiland adds that reliance on moral testimony of those in other social positions further has the practical value of respecting the moral experience of other agents with different social positions (cf. Wiland (2021), 100–104).²²

Apart from one’s social position, factors such as time constraints, a high mental load, the sheer complexity of an issue, etc. can all give good reasons to rely on moral testimony instead of deliberating on one’s own.²³ In other words, we need not assume that moral deference indicates something non-ideal about the *character* of the deferring person, but instead something non-ideal about the *environment* an individual finds themselves in—often caused not by themselves but by external constraints.

The common thread in these observations is that agents in non-ideal environments would limit themselves significantly if they decided to regularly forego moral testimony—because they lack access to certain epistemic resources bound up with a social position they do not occupy or accessible only if they were to invest time/energy that they currently do not have. There is, of course, a question as to whether these non-ideal environments should be accepted or whether they should be challenged, especially so by agents in privileged positions. However, we still need to provide the resources the agents in these environments need while *and in order to change* the status quo—and one such resource is reliance on moral testimony.

While we should ultimately strive to eliminate most of these non-ideal environments, there are other factors that might simply be acceptable non-ideal conditions. Morality is a vast domain that requires high degrees of specialisation: consider for example activists who specialise on one particular political issue or members of a social group strongly affected by another. These will be

²²Though Dular (2017) raises some problems for reliance on moral testimony under oppression—roughly, empirical evidence suggests that privileged (and oppressed) agents are more likely to listen to the testimony of other privileged agents.

²³Indeed, Kolers argues that deference in *political* matters is mandated regularly, simply because of the complexity of political issues. Not being able to defer would effectively immobilise the agent (cf. Kolers (2016), 77–81).

able to competently judge issues related to this field, giving guidance to others, while at the same time relying on guidance from others when it comes to a field they are unfamiliar with. Such a division of epistemic moral labour, however, does not strike me as inherently problematic—that is, unless we mistakenly assume that moral testimony itself is a problem.²⁴

This reveals another misconception of the debate around moral testimony. Using the label moral testimony suggests that morality is such a unified domain that one’s level of competence is uniform throughout the domain. Especially considering messy real-world cases, such an assumption of unity is surely mistaken. Even if you are competent in one domain of morality, you may be less competent in others. Often, then, agents deferring in ‘the moral domain’ will do so only on certain matters whereas they will contemplate on their own on others. Recognising this disunity in morality renders moral deference even less weird—most agents will defer every now and then, but certainly not on every moral matter. And if what I have argued is correct, such deference is often permissible—especially so if you do it for the right reasons.

2.5 Conclusion

In this chapter, I investigated how far-reaching pessimism about moral testimony is. I have argued that neither moral nor psychological sources support proper pessimism about moral testimony. Rather, a moderate optimism about moral testimony is appropriate: while there may be something regrettable about deferring in moral matters (a lack of understanding or a lack of motivational and affective states), this does not render moral testimony unusable. In comparing cases of non-ideal deferring agents with agents who try to figure out what to do on their own despite knowing that they could appeal to moral testimony and thereby improve their chances of getting it right, I have provided some reason to think that whatever is regrettable about moral testimony is relatively insignificant. This is not to say that one should always defer. If one is competent in a particular domain in morality, then deference to others is not the best way forward.

Recognising moral testimony as a legitimate resource in moral inquiry becomes all the more important once we recognise that morality is a vast domain with many domains that require specialisation. With limited time and resources, we will likely be unable to specialise in all domains we have to act in. If so, then relying on moral testimony will be an essential part of moral inquiry in these areas. Given the findings in this chapter, this should not trouble us. Deferring for the right reasons is not objectionable, even if it may come with some non-ideal consequences. Such consequences are what we would expect from the inquiry of non-ideal moral agents in non-ideal circumstances. In allowing for moral testimony as a legitimate resource, we, as non-ideal moral agents, can join

²⁴Some such divisions of labour *can* be problematic, and—given the inequalities in our societies—often are. For a very prevalent class of these cases, cf. Berenstain (2016).

together and try to foster a more cooperative understanding of moral inquiry.
There is no need to be all that pessimistic about moral testimony.

Chapter 3

I think I got it right

In this chapter, I present a novel argument for optimism about moral testimony, the view that it is permissible to rely on moral testimony when forming one's moral beliefs. The argument is simple: it is permissible to rely on one's own moral beliefs formed in the past even if the underlying reasons for the belief are no longer accessible to the agent. But there is no (relevant) difference between relying on one's past self and other agents—or if there is, it cuts the wrong way. Recognising that reliance on *intrapersonal* moral testimony is permissible should thus lead us to conclude that reliance on *interpersonal* moral testimony is likewise permissible. I explore some consequences for both optimism and pessimism about moral testimony following these considerations and argue that the present discussion offers a more nuanced understanding of what optimism about moral testimony amounts to.

3.1 Introduction

Every so often, I give some of my left-over money at the end of a month to charity. It does not matter here which charity it is, suffice it to say that it has been the same charity for about four years now. Back then, I made a conscious decision to give to said charity, thinking about the alternatives, weighing the reasons and coming to a conclusion. I must admit, however, that the charity I chose in the end was not a clear pick as it came down to a rather complex processes of weighing pros and cons. Nonetheless, I like to think that I made the right decision, basing it on the right reasons for why giving to that charity is the right thing to do. Yet, the process was complex enough that now, four years later, I could not recite the reasons for why the charity in question came out on top over other, reputable causes. Still, I continue to give to that same charity, without double-checking my decision each time. I give to the charity ‘merely’ based on the trust I have in my past self, its reasoning process and its eventual conclusion—but again, I think I got it right back then. What is more, I think that my giving to the charity is praiseworthy—I am deliberately doing the thing that is in fact right, and it is, in a relevant sense, no accident that I am doing the right thing.¹

Here are some questions that might come to mind: how do you know that you got it right back then? Why don’t you go and double-check? How could your giving to charity be praiseworthy, if you don’t even know anymore why it is the right thing to do? If you have these or related questions, I assume that you are a philosopher, probably even one invested in the recent literature around blame and praise, or moral testimony, and I will do my best to address these worries throughout this chapter. I would like to note, however, that I think the most pressing question for anyone who hasn’t been following the latest literature on these matters will rather be why I am formulating such an ordinary process in such a weird way: isn’t that what we all do, much of the time?

Well, yes—or so I think. And hoping not to make things worse, I must admit that the case of giving to charity is not the only one in my life where I simply trust my past self. My diet, my style of teaching, many of my decisions in relationships, and much more all depend on trust in my past self. Indeed, I rarely make conscious moral decisions relying on a reasoning process best understood exclusively in the present tense. It is much more common for me to make moral decisions relying at least partly on the fact that I know what I am about to do is right because I trust my reasoning process in the past. To illustrate, think of

¹You might worry that I am misdescribing my own mental state here. For couldn’t we hold that, *really*, there is a sense in which I am aware of the reasons even though they are not part of my occurrent beliefs? Perhaps there is some sense in which I *do* know the reasons for giving to the charity, even if I am not presently able to articulate them. This is ultimately an empirical question and so I find it difficult to argue against it on a conceptual level. There are two points I should like to mention. First, by setting the example up in the way I did, I was hoping to exclude such a reading: the reasoning process was complex, difficult, and it seems a stretch for me to say that I am, even if only in some loose sense, aware of the right-making reasons. Second, it certainly seems *possible* to conceive of an agent who lacks even the non-occurrent beliefs.

me in a restaurant deciding to go for a vegan choice over other options—I am most certainly *not* thinking about morality, let alone particular reasons for or against certain diets at the restaurant. I’m trying to find out what convention the restaurant is using to signify the vegan dishes and choose the tastiest option out of those. In fact, if a clever omnivore were to challenge me on my veganism in the restaurant, I might not even be able to directly respond to some of their arguments. Nonetheless, I don’t give up my veganism on the spot, knowing that I considered these very arguments in the past and found that they do not hold up to scrutiny.²

Again, I think that this behaviour is perfectly ordinary, and I don’t think it would be better for me to recount my reasoning process each time before making a moral decision. To be clear: it’s not that I am making ‘random’ moral decisions. I am just trusting my past self to have gone through the reasoning process well enough to ground my present decision. I like to think of this as trusting myself, relying on the moral testimony of my past self. And I don’t think there is any reason to think that this case is particularly special, there is nothing about *me* that makes trusting myself a good idea. And, more to the point, I don’t think there is a difference between relying on the moral testimony of my past self (let’s call this *intrapersonal* moral testimony) and relying on the moral testimony of others (let’s call this *interpersonal* testimony).

These two thoughts form the premises of the argument I put forth in this chapter. First, it is permissible to rely on your past self for moral testimony. Second, there is no (relevant) difference between relying on your past self and relying on someone else’s moral testimony. And so, it is permissible to rely on other people’s moral testimony.

The first premise falls out of the intuitions I sketched in the introductory example and I find it difficult to give a more principled argument in favour of it without entering the debate around (interpersonal) moral testimony directly. Thus, I take the intuitions at face value for the time being. However, after defending the second premise, I discuss the finding that results from the two premises, explaining in more detail why I think that reliance on intrapersonal moral testimony (and, given the argument, also interpersonal moral testimony) is legitimate. If my overall argument is successful, we have a novel argument in favour of optimism about moral testimony—the view that it is permissible to rely on moral testimony when forming one’s moral beliefs—and we have learned something about trusting ourselves along the way.

Connecting reliance on past reasoning processes of one’s own and reliance on the word of others is not a novel idea. In fact, Burge (1993) defends reliance on the word of others by drawing on the analogy to relying on one’s own findings

²This might seem at odds with my conciliationist leanings I outlined in chapter 1. However, recall that moral conciliationism is concerned with cases where (i) all evidence and arguments have been exchanged and (ii) we are facing a peer. If these conditions were met, my confidence in my past self would be much less secure. So, I am not defending a quasi-steadfast position here according to which we can always retain our beliefs based on the trust in our past self. I am merely claiming that there is nothing *intrinsically* wrong with relying on one’s past self—there might, of course, sometimes be reasons not to do so.

in the past. Burge argues that we are entitled to relying on both of these, where “entitlements are epistemic rights or warrants that *need not be understood by or even accessible to the subject*” (Burge (1993), 272–273, my emphasis). There is a sense in which I am following this basic idea, applying it directly to the contemporary debate around moral testimony.

3.2 The Parity Claim

In her 2012 defence of moral testimony, Sliwa discusses whether the apparent impossibility of assessing other agent’s reliability in moral matters is a reason to reject the legitimacy of reliance on moral testimony. Her response to this worry as brought up by McGrath (2009) is twofold. The more important part of the response, as Sliwa sees it, is that this impossibility extends not only to moral matters, but to other matters where we usually agree that reliance on testimony is perfectly legitimate—consider perceptual and mathematical testimony. There is, however, a second part of her response that I find to be equally consequential. If we are not able to assess the reliability of *other* agents in moral matters, Sliwa argues, “then it’s hard to see why it would be any better for an agent to rely on her *own* judgment [...]—she is no more in a position to check whether she herself got it right than she is in a position to check whether someone else got it right” (Sliwa (2012), 191, my emphasis).

If we are to reject the legitimacy of relying on moral testimony, or so Sliwa suggests, then we likewise should stop relying on ourself, since the two are “epistemically on a par” (Sliwa (2012), 191). Now, I don’t think we need to go as far—after all, it might be that reliance on moral testimony is epistemically worse off than making your own judgment, because you can base your judgment on morally relevant features while your reliance on moral testimony would not be so based.³ Nonetheless, I think that Sliwa has a point here: if we were to reject reliance on moral testimony by others as illegitimate, then so should *reliance on your past self* in moral matters, i.e. reliance on *intrapersonal* moral testimony, be rejected as illegitimate. It is these two that are epistemically on a par. Call this the parity claim, which just is the second premise of the argument put forth in this chapter: there is no (relevant) difference between relying on your past self and relying on someone else’s moral testimony.

Of course, the literature on moral testimony focuses on cases of *interpersonal* moral testimony, testimony coming from other agents to us, which—if the pessimist about moral testimony is right—we should usually reject. So, if the pessimist wants to uphold that position and simultaneously hold that it is permissible to rely on *intrapersonal* moral testimony, they have to point out a relevant difference between the two. In this section, I go through some possible differences and reject them as making a (relevant) difference.

What is a relevant difference for the purposes of this discussion? Plausibly, it is one that also makes a difference in evaluation, i.e. a difference the pes-

³I discuss the immediate relevance of Sliwa’s finding regarding the question whom to trust further in chapter 5.

simist could point to in order to explain why reliance on interpersonal moral testimony is illegitimate while reliance on intrapersonal moral testimony is not. And the best place to look for such differences just is in the motivations pessimists about moral testimony provide for the alleged illegitimacy of reliance on (interpersonal) moral testimony. If there *is* a relevant difference, it better be found in the motivations pessimists have brought forward—for these should then *not* motivate the same concerns about intrapersonal moral testimony.

To allow for a more pointed discussion, I group the various motivations for pessimism about moral testimony, hoping to nonetheless give each individual proposal sufficient space. Following Hopkins (2007)’ distinction between unavailability and unusability accounts, all of the proposals I consider fall under the heading of unusability accounts. That is, all of these pessimists I discuss agree, as is now commonplace, that moral testimony does make moral knowledge available (as opposed to the unavailability accounts), but that such knowledge is, for some reason or other, illegitimate to use.⁴

Ultimately, I conclude that there is no such relevant difference. From the perspective of the committed pessimist, I think, intrapersonal moral testimony is just as bad as interpersonal moral testimony.

3.2.1 Moral Understanding

Perhaps the most influential motivation for pessimism about moral testimony is due to Hills (2009)’s suggestion that testimony-based moral knowledge is unusable because moral testimony fails to convey moral understanding. However, or so the argument goes, moral understanding is particularly valuable because it allows us to perform actions of moral worth, doing the right thing *for the right reasons*—an independently valuable aim (cf. Hills (2009), 108–119). In a similar vein, Nickel (2001) has argued that “morality requires one to *act* from an understanding of moral claims, and therefore to *have* an understanding of moral claims that are relevant to action” (Nickel (2001), 257, author’s emphasis). Again, reliance on moral testimony fails to convey such understanding, thus leaving the resulting belief deficient with respect to its justificatory basis. A third suggestion originating broadly from the same camp is due to Hopkins (2007), although it is less clear whether he takes it to succeed as a proper reason to be pessimistic about moral testimony. In discussing which route to pessimism is most promising he suggests that “having the right to a moral belief requires one to grasp the *moral* grounds for it” (Hopkins (2007), 630, author’s emphasis).

Can moral understanding make a relevant difference in cases of reliance on your past self and others? After all, by the setup of the case, there is a point in time at which I *did* have that moral understanding, and so perhaps I still

⁴Most of the time, unavailability concerns are only brought up to reject their plausibility in favour of unusability concerns (cf. e.g. Hopkins’s own discussion thereof in Hopkins (2007), 615–626). A notable exception is McGrath (2009), who argues that we cannot assess the reliability of moral testifiers, thus making moral knowledge unavailable in the first place. I hope the brief discussion of Sliwa’s response above is sufficient to set this worry aside—after all, Sliwa’s point *motivates* the parity claim.

gained the relevant achievement—a difference to the agent who simply defers to someone else’s testimony without prior understanding. The proposed difference between cases of intra- and interpersonal moral testimony would thus be that moral understanding need not be present *now*, as it were—it suffices if it was present at some point. However, the proposed difference doesn’t just cut across intra- and interpersonal moral testimony. For consider the following two cases of interpersonal moral testimony:

No understanding in the past: Eleanor has always enjoyed eating meat but has recently realized that it raises some moral issues. Rather than thinking further about these, however, she talks to a friend, who tells her that eating meat is wrong. Eleanor knows that her friend is normally trustworthy and reliable, so she believes her and accepts that eating meat is wrong. (Hills (2009), 94)

Understanding in the past: Eleanora used to be an ethical vegetarian in her youth, understanding why that was morally required of her. However, growing older she cared less about these matters and started enjoying eating meat again. Recently, however, she remembered it raises some moral issues. Rather than thinking further about these, however, she talks to a friend, who tells her that eating meat is wrong. Eleanora knows that her friend is normally trustworthy and reliable, so she believes her and accepts that eating meat is wrong. (Adapted from Hills (2009), 94)

If moral understanding at some point in the past is enough to make the difference, then Eleanor and Eleanora are to be evaluated differently. But that seems absurd, especially from the pessimist’s perspective. Both Eleanor *and* Eleanora should think about the issue themselves, or so the pessimist is likely to uphold, and so both shouldn’t form their beliefs relying on their friend’s testimony. For the understanding-styled pessimist, it simply shouldn’t matter what your past beliefs were—you should not form your beliefs basing them on something other than (the right) reasons, its (actual) moral grounds.

There is an obvious reply on behalf of the pessimist to try and uphold their position. They might agree that Eleanora, like Eleanor, fails to satisfy the relevant criteria, because her belief now is in no way *connected* to her past reasoning process (to see why that is, we can consider an alternative case similar to the one of Eleanor—even without a prior reasoning process, the resulting belief remains the same due to her deference). The more precise way to formulate the difference would thus be to hold that not only do you need to have understanding at some point in the past, that understanding has to be connected to your current belief. Thus, all cases of interpersonal moral testimony would fail the test whereas reliance on intrapersonal moral testimony would be legitimate. The problem now becomes to spell out what exactly such a connection is supposed to amount to—other than having moral understanding *now* in the first place.

But this reply does not help the pessimist about moral testimony who is motivated by concerns around moral understanding anyway. For if moral understanding is valuable because it enables us to perform morally worthy actions,

to act from moral understanding, etc., then not only do both agents above, Eleanor and Eleanora, fail the test, the same goes for the introductory example. Not being able to recount the interplay of the reasons delivering the verdict that I should give to the charity in question, I am at best motivated to do the right thing, I am most certainly not motivated by the actual right-making features of my action.

If testimony-based moral beliefs are bad because they fail to convey moral understanding, I submit, both reliance on intra- and interpersonal moral testimony is illegitimate. Moral understanding fails to make a (relevant) difference.

3.2.2 Conative Attitudes

Some pessimists about moral testimony have suggested an alternative to this explanation, arguing that moral understanding is focused too much on cognitive elements of belief, while it actually is its conative elements are of crucial importance in the moral domain. Yet, there is some continuation in the tradition of the understanding-styled pessimist: Callahan (2018), for example, considers such a proposal as “reconceiving understanding as a richer state than it is commonly thought to be, comprising affective and motivational engagement with reasons” (Callahan (2018), 438). Such affective and motivational components, however, fail to come along with moral testimony. Similarly, Fletcher (2016) draws attention to the conative elements of moral belief, arguing that it “is impossible to form desire-like moral sentiments (states such as e.g. anger, blame, guilt and resentment) on the basis of pure, direct, testimony” (Fletcher (2016), 60).

Does this extension help the pessimist to identify a difference between intra- and interpersonal moral testimony? That is, assuming that it is not possible to form conative attitudes on the basis of interpersonal moral testimony, is the same possible on the basis of intrapersonal moral testimony?

Both questions are ultimately empirical questions. For the purposes of discussing the parity claim, however, we need not settle these questions themselves, we merely need to settle whether they could receive a *different* answer. And it is hard to see why they would. The reason why Callahan and Fletcher think that testimony-based moral beliefs come without their usual conative elements is the disconnect to insights into moral reality (here, we can most clearly see why these accounts can be thought of as an extension of the moral understanding account discussed previously). But, by hypothesis, this connection is also missing in the cases of intrapersonal moral testimony we are concerned with here.

Upholding that intrapersonal moral testimony *can* provide the conative elements to the resulting moral belief would either be misdescribing the cases we are interested in (by artificially creating such a connection) or else undercut the initial argument for the lack of conative elements in moral beliefs resulting from reliance on interpersonal moral testimony (by giving up on the importance of the connection). And that is precisely what we would expect given the parity claim: there is no (relevant) difference between the two. Indeed, in a 2020 article, Doyle argues on this very basis that *both* reliance on moral testimony

by others and “remembering what is right”, as her paper is aptly titled, are illegitimate (cf. Doyle (2020)).⁵

The common reason why neither of the moral understanding-based motivations can identify a relevant difference between intra- and interpersonal moral testimony, I take it, is due to their synchronic nature. If all we focus on is the present state of the agent, there simply is no difference between intra- and interpersonal moral testimony. In introducing cases of reliance on one’s own past self, the key change is the newly introduced diachronic nature. Perhaps, then, motivations more intimately connected with such diachronic considerations might be of help.

3.2.3 Autonomy

One such candidate could be autonomy—as it is a general capacity, rather than a mere state of mind. In the 88th volume of the *Aristotelian Society Supplementary*, Lillehammer (2014) and Crisp (2014) both discuss the value of moral testimony from the perspective of autonomy—Lillehammer arguing that there is no distinct problem arising here while Crisp (2014) holds that there is. In setting up the discussion, they draw on Wolff (1970)’s conception of an autonomous being:

He may listen to the advice of others, but he makes it his own by determining for himself whether it is good advice. [...] He may do what another tells him, but not because he has been told to do it. [...] For the autonomous man there is no such thing, strictly speaking, as a *command*. (Wolff (1970), 13–15, author’s emphasis).⁶

On Wolff’s understanding of autonomy, reliance on moral testimony restricts one’s autonomy and is thus deficient in this way. Lillehammer finds this suggestion lacking, as it is implausibly strong—one’s decision to defer is, after all, limited to certain occasions and could itself be the result of autonomous deliberation. One’s capacity to be autonomous would thus not be undermined. Even if there is some value to autonomy, it is therefore unclear whether it undermines the legitimacy of relying on moral testimony.

Considering this response, Crisp reformulates the initial worry to encompass Lillehammer’s criticism. Crisp argues that while in relying on moral testimony, the agent does not hand over their capacity for autonomy, “she fails to *exercise*

⁵Notably, Doyle engages neither with Fletcher’s nor with Callahan’s work. Nonetheless, her assumption of at least a weak version of motivational internalism as well as her concern with emotional propensities suggests that her account is closely aligned with the accounts I discuss in this section (cf. Doyle (2020), 57–59).

⁶Of course, Wolff thinks that it can sometimes (indeed, often) be reasonable to give up one’s autonomy. Yet, the alleged accompanying costs of ‘losing’ autonomy are in an interesting tension with the thought expressed in the epigraph of this dissertation—Freire’s insistence that a “real humanist can be identified more by his *trust* in the people, which engages him in their struggle, than by a thousand actions in their favor without that trust.” (Freire (2000), 60, my emphasis). I doubt that Freire would consider the Wolff’s worries as legitimate worries about autonomy. They would be appropriate only insofar that a distorted (and problematic) notion of the *oppressor’s* autonomy is concerned.

her autonomy *on this occasion*” (Crisp (2014), 137, my emphasis), and that is all that matters. This response, however, reveals that *this* worry about autonomy is ultimately *not* diachronic, but synchronic after all. While we need not discuss whether autonomy is ultimately successful as a motivation for pessimism, we can now see that the particular understanding of autonomy as employed here does not threaten the parity claim precisely because it is not the capacity, but the exercise that is under threat in relying on moral testimony: if the problem with reliance on moral testimony is its failure to exercise one’s autonomy on some particular occasion, it is difficult to see why this same problem would not hold for intrapersonal moral testimony (even if the agent as a whole retains their capacity for autonomy, as it were).

Another problem Lillehammer brings up for autonomy as a motivation for pessimism about moral testimony is that while the deferring agent does not presently grasp the moral grounds for their testimony-based moral belief, they are not in principle inaccessible (cf. Lillehammer (2014), 122). However, because Crisp thinks that it is not the capacity but the exercise of autonomy that is of relevance here, he responds by noting that “merely potential or hypothetical grasp of reasons is insufficient for autonomy” (Crisp (2014), 139). Whether or not this is successful as a defence, it is again revealing for our purposes here. If what matters is not just the potential grasp of reasons, but the actual grasp thereof, then, again, intrapersonal moral testimony does not fare better than interpersonal testimony.

I take it that these autonomy-based motivations fail to identify a difference between intra- and interpersonal moral testimony due to their rather synchronic understanding of which aspect of autonomy is of relevance to the pessimist’s project. Once more, we would need a more diachronic concept to uncover a potential difference. However, there is one such final contender: authenticity.

3.2.4 Authenticity

There are at least two accounts in this vicinity: Howell (2014) suggestions of a problematic character in the deferring agent and Mogensen (2017) more explicit appeal to authenticity. There are some differences between the two accounts, but they do not make a difference for the discussion at hand.⁷

Howell argues that reliance on moral testimony indicates a deficient moral character because such testimony-based moral beliefs fail to be subjectively integrated in one’s character (cf. Howell (2014), 408). One’s character is plausibly not just recorded at some point in time, but over time—thus incorporating the newly introduced diachronic element. Similarly, Mogensen motivates pessimism about moral testimony drawing on the value of authenticity.⁸ An action

⁷The key differences (and simultaneously the reasons to prefer his account) Mogensen outlines are that Howell’s account cannot explain why agnosticism sometimes seems preferable to deference (I myself am actually doubtful of this intuition) and that it is not just the mere lack of conative attitudes that is important but that the lack is due to the agent’s deference (cf. Mogensen (2017), 273–274).

⁸Actually, Mogensen is not explicitly endorsing pessimism about moral testimony. In detailing his understanding of authenticity, he raises some worries about the value attached to

performed on the basis of moral testimony, Mogensen argues, cannot be an authentic moral action—but authenticity is important when it comes to morality (cf. Mogensen (2017), 280).

Here is Mogensen’s characterisation of authenticity:

Authenticity is achieved when the beliefs and motives that guide a person are expressive of her *true self*. Inauthenticity arises when these states fail to give expression to who we really are. [...] To be authentic, the beliefs which guide us through life must give expression to the *true self*. This seems to require that we should decide moral questions on our own terms, so far as we can, so that our own moral sensibility is manifest in the values and ideals by which we live. (Mogensen (2017), 276–277, my emphasis)

The ‘true self’ Mogensen refers to is in some ways similar to Howell’s highlighting of the importance of character, although it is perhaps the more substantive approach. Either way, the authenticity-styled theorist à la Howell or Mogensen seems to be able to reject the parity claim: insofar as your past self and your present self are expressive of your true self/character, reliance on such intrapersonal moral testimony is permissible. This seems to mark a difference to regular cases of interpersonal testimony and would also allow us to reject putatively impermissible instances of intrapersonal moral testimony. For suppose your outlook on life changes dramatically. Even if we think that reliance on your past self is in general permissible, it does not seem that we should rely on such a dramatically different past self—and the authenticity theorist has a straightforward explanation for this phenomenon: your true self/character changed.⁹

The authenticity-styled pessimist thus *does* seem to have a way to mark a difference between intra- and interpersonal moral testimony, allowing them to explain why reliance on the one (your past self) is permissible while reliance on the other (other people) is not. I worry, however, that the difference we might thus draw is ultimately ill-motivated.

Why is it permissible, on the suggested response of the authenticity theorist, to rely on your past self? Plausibly, because something along the lines of the following counterfactual holds: if you were to do the reasoning all over again now, you would come to the same conclusion—after all, your past and present self are both expressive of your true self/character. Conveniently, as we saw above, such a counterfactual also explains why reliance on your past self in cases of dramatic changes in your outlook is not permissible: if my outlook changed drastically enough, there is no such assurance that I would come to the

it—but thinks that the question whether authenticity really should be an ethical ideal needs to be settled elsewhere (cf. Mogensen (2017), 276, 281). Nonetheless, he at least suggests that authenticity is a very plausible candidate for an ethical ideal, thus supporting pessimism about moral testimony (cf. Mogensen (2017), 277–278).

⁹This also helps in considering cases where beliefs change over the course of one’s lifetime: which of my past selves, so to speak, should I trust? Plausibly, the beliefs of the past self that is most congruent with one’s current self—and this needn’t always be the most ‘recent’ past self.

same conclusion, and so relying on my past self in this way would leave me with an ultimately inauthentic belief.

While this does mark *a* difference to interpersonal moral testimony, I think the difference does not cut the right way. We don't usually rely on other people's moral testimony because we believe that we would come to the same conclusion as they would. In all plausible instances of relying on someone else's moral testimony, we seek out their testimony precisely because we think that they are *more likely* to get things right, that they have access to some epistemic resources we lack access to.¹⁰ And so rather than marking the difference in the right way, it strikes me that these considerations pose a considerable problem to the authenticity theorist, a problem that is similar to the problem I outlined in chapter 2 for Hills' proposal: why should the value of authenticity be so high that it outweighs getting things right in the first place?

There are at least two responses available to the authenticity-styled pessimist about moral testimony here. First, they might accept the argument so far, but dig their heels in and hold that the value of authenticity just is very high. I think this route is not particularly promising, and—again, it is for similar reasons to the ones I considered when discussing the possible response on behalf of Hills and Howell in chapter 2: whatever right-making feature the pessimist comes up with, it is at best derivative from the primary aim of moral inquiry—getting things right. And the same thought applies here—if authenticity has any value in moral inquiry at all, it is difficult to see why it should ever trump the primary aim of moral inquiry—getting things right. Mogensen's own hesitance to endorse the value of authenticity all the way might be explained in a similar fashion.

But there is an alternative response available to the authenticity theorist. For who would be better placed than me when it comes to assessing what matters (morally) than myself? This response casts doubt on the idea that there *could* be anyone else better placed to respond to moral matters from the perspective of the values I uphold. Given the centrality of one's own values, plausibly deriving from one's true self/character, I take it that this response is much more in the spirit of the authenticity theorist. The idea, then, is this: I value certain things—e.g. friendship, charity, etc. Given that I am the only one of whom I can be certain to assess matters from the perspective of exactly these values (as well as how I understand and weigh them), there simply is no one else who I could confidently judge to be better placed than me when it comes to responding 'in my way' to moral matters.¹¹

¹⁰Sometimes, the relevance of moral testimony is 'explained away' by pointing out that in relying on someone else's moral testimony, we really only seek out people whose moral values are sufficiently similar to ours so that we end up deferring only on some empirical questions that they might have more information about. While those cases certainly exist and are not instances of 'real' moral testimony, they do not exhaust the potential instances of moral testimony. It is possible to defer to someone on moral matters even if they have different moral values to you—and I take it that those are the cases most optimists (and pessimists) about moral testimony are concerned with.

¹¹Note that this is *not* the same worry as a more general worry regarding identifying other agent's moral reliability expressed by McGrath (2009). She argues that it is the lack of

There is a sense in which this seems right—who, after all, could know what you actually value better than yourself? (Well, your therapist perhaps, but that’s besides the point.) The answer to this question depends at least in part on how you conceive of the things you value. Consider friendship. On one way of understanding things, you might value certain things, say, going to great lengths to help someone out, intimate ways of relating, etc., and hold that these things are what you take friendship to be. If so, then it does indeed seem that you are the sole authority when it comes to judging moral matters bound up with friendship. Plausibly, no one else values these things exactly as you do. However, there is an alternative way of understanding matters: again, you value friendship, and understand friendship to be comprised of those same things as before. This time, however, you consider yourself to be ‘open to correction’. That is, you are open to the possibility of being wrong—perhaps going to great lengths to help someone out turns out to be insignificant when it comes to friendship. On this way of understanding things, there could be someone else better placed than you to understand what friendship means.¹²

So, how *should* we value the things we value? This question often receives an all-or-nothing answer, favouring either one of the readings outlined above. As I argue in the next section, this framing is unfortunate because it creates somewhat false opposites. For now, however, we merely need to be concerned with an existential claim: is it ever permissible to value in the second way? And it seems indeed undesirable to negate this existential claim, for otherwise it would secure the result the authenticity theorist needs solely because our true self would be effectively incorrigible by external forces: if I care only about the things that I actually value, as narrowly described as possible, and not because they are an instance of some more general level of description, then it is *only* through changes in my valuing-structure that I can change the objects of my valuing.

But this is too strong a requirement for authenticity: evidently, my true self/character can also authentically value, say, friendship, if I leave myself open to not having the final say about what would realise friendship, about what friendship really is, if you will (and the same goes for moral rightness). Suppose you change your mind about what friendship really means after discussing the matter with someone else. If we allow for the possibility of valuing also at a more general level of description, then we can easily describe the situation as one in which you were wrong about what friendship really means (and what it really means to you) before, and are now getting closer to its ‘true’ meaning. If, however, we don’t allow for such a valuing structure, then your prior understanding (and the corresponding valuing-structure) of friendship wasn’t misguided at all, it’s just that you valued something different before, and value

an independent check in the moral domain means that we cannot assess the reliability of a putative testimony-giver. More on this in chapter 5.

¹²This discussion might remind the reader of a debate around motivation by moral rightness *de re* and moral rightness *de dicto*. I agree with Johnson King (2022) that these labels have not been the most helpful ones in advancing the discussion for reasons I go into in the next section, and so I do without them for now.

something else now. Properly speaking, it is not your understanding of friendship that changed, it's just that some of your values changed, that's all. But the underlying psychology here seems odd: weren't you aiming at something when you valued friendship—and doesn't aiming at something just entail the possibility of missing the target?

So, unless the authenticity theorist wants to suggest such an unappealing characterisation of what it means for our true selves/characters to authentically value something, they should leave themselves open to learning from others. And if so, then the fact that intrapersonal moral testimony 'originates' in our true self whereas moral interpersonal testimony does not, cannot make the difference between the two. Our true self/character can be authentically involved *both* in intra- and interpersonal moral testimony.¹³

Let's take stock. There is an understanding of authenticity resulting in a relevant difference between intra- and interpersonal moral testimony, thus invalidating the parity claim. I have argued that this understanding of authenticity is implausible—and I don't think either Howell or Mogensen would want to commit to it. The alternative understanding of authenticity, however, does not invalidate the parity claim. If we understand authenticity as being true to ourselves where this allows for the possibility of error, motivations surrounding authenticity do not make a relevant difference between intra- and interpersonal moral testimony. As we have seen above, the same is true for motivations surrounding moral understanding, conative attitudes and autonomy, none of them support a relevant difference—the parity claim holds.

I suspect that at this point the pessimist might be inclined to switch gears. The discussion in this section, they might uphold, does indeed show that the cases of intrapersonal moral testimony I am interested in are not relevantly different to usual cases of interpersonal moral testimony. But that does not show that both kinds of reliance are permissible. Rather, it raises doubts as to whether reliance on your own past moral beliefs (in the way outlined) is permissible in the first place.

3.3 Trusting Yourself and Trusting Others

In this section, I try to understand why it might have seemed intuitively plausible to think that reliance on intrapersonal testimony is legitimate, even for pessimists about (interpersonal) moral testimony. Given the parity claim, however, this then leads me to arguing that reliance interpersonal moral testimony is legitimate for essentially the same reasons.

Insofar as intra- and interpersonal moral testimony are alike, it would be tedious to go over the arguments optimists and pessimists have put forth in

¹³Note though that this gives us a different result compared to the earlier motivations for pessimism, where both intra- and interpersonal moral testimony are equally *bad*. In the broader context, my line of reasoning here suggests that authenticity concerns support pessimism about moral testimony only under a certain, implausible, conception of authenticity. On the more plausible reading available, authenticity fails as source for pessimism about moral testimony in the first place—both intra- and interpersonal moral testimony are equally good.

favour or against reliance on interpersonal moral testimony. It doesn't strike me as likely that either side is going to find the other's argument more convincing in the present case. Instead, I want to suggest that many of the core elements found in the pessimist's motivations for rejecting interpersonal moral testimony are in fact not threatened by *intrapersonal* moral testimony. (And, given the parity claim, the same ultimately holds for *interpersonal* moral testimony.) This needs some explanation. Of course, I am not claiming that the motivations in the way as brought forward by the pessimist do not threaten the legitimacy of interpersonal moral testimony at all. However, I think investigating those motivations and the reasons for them in closer detail shows that we needn't be all that worried about (interpersonal) moral testimony after all—even if we care about the various values pessimists about moral testimony have brought up. In fact, I think many of the motivations are compatible with reliance on intra- and thus interpersonal moral testimony.

Before going down this road, I should mention the alternative way the pessimist about moral testimony has if they accept the parity claim: they can of course hold that *both* reliance on intra- and interpersonal moral testimony is illegitimate (and in the same way). While this is possible, it still undercuts the intuitive case for pessimism about moral testimony. Where we once had the intuitive “Isn't there something off about relying on other people's word for your own moral beliefs?” that we then sought to explain drawing on a lack of moral understanding, conative attitudes, autonomy or authenticity respectively, we are now left with the much less intuitive question offered directly by the explanations themselves: “Isn't there something off about moral beliefs that are lacking with respect to moral understanding, conative attitudes, autonomy, or authenticity?”. Well, maybe—but it is at least not immediately intuitive that this should be so. At the very least then, the parity claim shows that pessimism about moral testimony is a more expansive position than previously thought.

Perhaps this is a bullet some pessimists about moral testimony are willing to bite. After all, we shouldn't reject moral testimony because of the intuition, but because of the underlying reasons. However, even by the pessimist's own light, excluding reliance on moral testimony by one's past self comes at a cost. And this is where the attempts to mark a difference between intra- and interpersonal moral testimony from the previous section come back in. The dialectic is this: the pessimist initially sought to identify a difference between the two, in order to keep intra-, but reject interpersonal moral testimony. I argued that no such difference is to be found. The pessimist then takes this to be evidence that both kinds of testimony are equally bad. I take it to be evidence that both are equally good.

3.3.1 No Ulterior Motives

In chapter 2, I argued that (too) much of the debate around moral testimony is centred around agents who defer to others mostly because of ulterior motives. They are often characterised as lazy agents, deferring rather arbitrarily, often to shirk moral responsibility (cf. Sliwa (2012), 188–189, for a criticism of some of

these depictions). Though unfortunate, it is of course possible that some agents rely on moral testimony in this rather irresponsible way. For cases of reliance on *intrapersonal* moral testimony, however, such readings are almost always excluded: the agent relying on their past self cannot be motivated by any of these ulterior motives. After all, they *did* figure out things for themselves—it’s just that the past tense matters in their case.¹⁴

The reason even pessimists about moral testimony should see the appeal of the legitimacy of reliance on intrapersonal moral testimony is in part because a lot of the motivations for accepting it stem from the resistance to the parity claim I discussed in the previous section. Isn’t it right that in relying on intrapersonal testimony, the agent is trying to get at the right grounding for their moral decision? After all, they figured things out for themselves, even if they aren’t able to recall the reasons. Isn’t it right that they are appropriately involved in their moral decision? For they did care about the matter enough to think about it for themselves. Isn’t it right that the agent exercises their autonomy in deciding to rely on their past self instead of doing the reasoning process all over again? And isn’t it right that the agent wants to stay true to themselves, act authentically, and *therefore* decides to rely on their past self?

As we have seen, pessimism about moral testimony is often motivated by drawing on a missing connection between the resulting moral belief and the underlying reasoning process. This distance between the belief and the underlying reasons, however, shrinks once we consider the case of *intrapersonal* moral testimony. For it is of course no accident that I give to the correct charity (if I do) or that my diet is ethical (if it is). And, more importantly, the non-accidentality here is indeed connected to the right-making features of the act in question, because I *did* think about the matter in an adequate way. It is just that, at present, I am not able to recall said reasoning. So while my correct belief/action is not directly connected to the right-making features, a connection is nonetheless there, and it is strong enough to ground the relevant non-accidentality.

One way of explaining this connection would be to draw on the notion of *planning* as developed in Bratman (1987)’s account of intentions. The driving force behind his account is that we, as limited rational agents, need plans to direct our future selves. For Bratman, some of our intentions are in effect policy-based, and it seems likely that many of our moral convictions will be such policies.¹⁵ Crucially, such a policy does not require constant awareness of the reasons for the policy, precisely because we need not reconsider the policy every time we put it into action (cf. Bratman (1987), 91). Indeed, if we deliberately decide not to reconsider our policy, such non-reconsideration is not

¹⁴You might be tempted to thus locate the difference between intra- and interpersonal testimony in an ulterior motive such as shirking responsibility. However, this would require such a shirking of responsibility to be *successful* which in turn would require further argument. In chapter 6, I discuss the question of responsibility in reliance on other people’s moral testimony and argue that one does not give up responsibility for one’s actions by deferring.

¹⁵Indeed, some expressivists like Gibbard (2003) propose that this is the best way to understand moral conviction—developing a notion of *to-be-doneness*.

reason-preserving, but *reason-changing* (cf. Bratman (1987), 63). Despite the changes in our motivating reasons, however, the extended deliberation plans allow for ensures “that the reasons and deliberation that really are responsible [...] are included within the scope of our assessment” (Bratman (1987), 79). In short, such a notion of plans allows us to develop the right kind of connection between my intrapersonal testimony-based action and its right-making features.

But note that this non-accidentality of the connection is not contingent upon formulating my action (giving to charity, keeping my diet) as a plan—one-off examples secure the result just the same: consider, for example, someone who gave to a particular charity once in the past and, after receiving a substantial inheritance, decides to give to charity again. Without ever having planned on that, they could just as well rely on their past self for the decision as I can in the introductory example to this chapter.

I mention this because it might seem tempting to think that whether an action was planned or not makes a difference, for it might seem natural to think that in acting on a plan, I can still act for the reasons that led to my formulation of the plan, even if they are no longer occurrent to me (note that Bratman (1987)’s suggestion of reason-changing non-reconsideration also seems to allow for cases where this is not the case). Spelling this objection out would require a theory of what it means to act for a reason, but I don’t think we need such a theory to see that a difference between acting on the basis of a plan where one forgot the reasons for the plan and acting on the basis of ‘mere’ intrapersonal testimony without a plan where one forgot the reasons for the prior judgment is difficult to draw. If we have a rather internalist understanding of what it means to act for a reason, both actions will fail to qualify, but once we go more externalist, both planned action and merely remembrance-based actions will count as acting for the (no longer occurrent) reasons/right-making features.¹⁶ And, as I will suggest below, the same goes for moral advice by others.

The focus on the right-making features is suggested by all the motivations we considered for pessimism about moral testimony. Moral understanding as well as appropriate conative attitudes plausibly derive from awareness of the right-making features, exercising one’s autonomy just is basing one’s action on one’s reasons (here, the right-making features) and acting authentically means acting in accordance with those identified reasons.¹⁷

What I want to propose is that the same holds for cases of interpersonal moral testimony, even if it might be less obvious (because I was *never* aware of the right-making features). First, note that reliance on someone else’s moral testimony is of course only permissible when you have good reasons for taking your interlocutor to be reliable with respect to the proposition in question. And,

¹⁶The only way a plan *could* make a difference is in its securing a higher reliability of your past judgment. If I plan to give to a certain charity regularly, this might—other things being equal—make it more likely that I will correctly remember and act on my decision than if I merely remember that I decided to give to said charity in the past.

¹⁷Lord (2018) considers a lack of acquaintance with the ethical facts itself as the problematic feature of moral testimony. In a sense, then, many pessimists about moral testimony (and, in particular those we surveyed here) can be understood as providing an explanation as to why this feature is problematic.

at least in all usual cases, that will be precisely *because* you take them to be well equipped (better equipped than yourself at any rate) to track the right-making features in question. Therefore, when all goes well, your believing or acting on the basis of interpersonal moral testimony, as in the intrapersonal case, is non-accidentally connected to the actual right-making features of the belief or act in question.¹⁸

When I defer well to someone else, it is simply no accident that I end up doing the right thing. And if we should care about deliberately doing the right thing, as Johnson King (2020a) argues all sides of the debate can agree on, then moral testimony is—under the right circumstances—not problematic at all: “someone can deliberately do the right thing even if she is not at all confident that her act is morally right, if she is trying to act rightly and succeeding” (Johnson King (2020a), 203–204). And this is just a description of a typical case of reliance on interpersonal moral testimony—and, similarly, a description of many cases of reliance on intrapersonal moral testimony.

So far, my suggestion is to reject the detachedness of interpersonal moral testimony from the moral grounds for action, in particular by drawing on considerations resulting from the analysis of intrapersonal moral testimony. But the pessimist might hold that the connection simply is not direct enough, so let me explore this suggestion further in the remainder of this chapter.

3.3.2 A More Nuanced Optimism

In chapter 2, I discussed briefly how the debate around moral testimony is sometimes framed around the question of proper moral motivations. Drawing on Smith (1994), it is suggested that we can use the *de re/de dicto* distinction to illuminate these motivations as follows.¹⁹ On the one hand, an agent might be motivated by an explicit concern with moral rightness itself (supposedly motivation by moral rightness *de dicto*). On the other hand, an agent might be motivated by the various right-making features (supposedly motivation by moral rightness *de re*).

This framing is then employed in the debate around moral testimony by arguing that reliance on moral testimony, in virtue of failing to deliver awareness of the right-making features, cannot be motivation by moral rightness *de re*, and, since it requires an explicit concern with acting rightly, has to require motivation by moral rightness *de dicto*. While this result is sometimes disputed, most notably by Wiland (2021), it is often readily-accepted also among optimists about moral testimony who then assume a position that reduces the importance of the awareness of moral features in favour of explicit concerns with acting rightly. I think this response is a mistake because the framing the debate

¹⁸Indeed, Roth (2017) argues that reasons can sometimes be interpersonally transmitted, i.e. it is possible to act for a ‘reason of another’ (cf. Roth (2017)). He acknowledges that this position is controversial, and we need not go all the way to support the point I want to press here.

¹⁹Note that Smith himself does *not* endorse the application in this way.

receives is inaccurate and, as a consequence thereof, there is a more nuanced understanding of optimism about moral testimony available.

The criticism of the way the *de re/de dicto* distinction has been applied in the debate around moral motivation is due to Johnson King (2022). She points out that this framing creates a wrong dichotomy between certain kinds of motivations and obscures other kinds of motivations. In particular, the contemporary framing allows for only two kinds of concerns—an explicit concern with acting rightly and a concern with right-making features. This framing obscures another concern: a concern with acting rightly under *another* description.

It is the *de re*-qualifier in particular that has been applied misleadingly. The usage in the literature suggests that motivation by rightness *de re* means that the agent in question is motivated by the right-making features of an act. However, the qualifier really is supposed “to signal that the object of an agent’s attitude is the very same entity to which our description refers, though she herself may consider it under a different description. [...] But right-making features are not the property of moral rightness considered under another description” (Johnson King (2022), 259). The correct application of the *de re* qualifier would thus fall in line with the previously obscured kind of motivation: a concern with acting rightly under another description. To see why this is relevant, Johnson King describes some of these motivations, holding that they are commonplace for many moral agents: striking the right balance, finding out what *really* matters, etc. (cf. Johnson King (2022), 268).

Johnson King’s proposal now is that these previously obscured motivations (which are properly described as motivation by moral rightness *de re*) are unproblematic motivations even if they include a concern with moral rightness—albeit under another description. Drawing on this newly identified kind of motivation, Johnson King develops an argument against the putative problematicness of being motivated by moral rightness *de dicto*, i.e. of explicit concerns with acting rightly. If the newly identified motivations are fine, then it cannot be the *object* of this motivation that is problematic. Instead, the problem must arise from the *content* of motivation by moral rightness *de dicto*. But once we acknowledge that “there is nothing wrong with being motivated to act rightly under another description, then why would there be something wrong with being motivated to act rightly *under this very description*” (Johnson King (2022), 255, my emphasis)?

With this newly opened up theoretical space, the optimist about moral testimony does not need to go in for caring only about explicit concerns with moral rightness (although such concerns are, following the argument above, equally unproblematic). And I suggest that we can see this in action when considering intra- and interpersonal moral testimony. I outlined earlier that reliance on intrapersonal moral testimony seems in part less problematic because it is reliance on moral considerations of one’s own. As such, there is no need to construe your reliance on your past self as being motivated by an explicit concern with acting rightly (under that description). You might equally well be concerned with striking the right balance (again), reminding yourself of what really matters, etc. Of course, you *could* be motivated by an explicit concerns with

moral rightness, but I suspect that the attraction of many cases of intrapersonal moral testimony stems from the fact that no such explicit concern is necessary to get the cases going. This is why the renewed charge of moral fetishism as brought forward by e.g. Doyle does not succeed in delegitimising reliance on intrapersonal moral testimony (cf. Doyle (2020), 58).

For interpersonal moral testimony, it is at first glance more difficult to see how an explicit concern with moral rightness could be missing—partly for the reasons we saw in the discussion of authenticity as a potential parity breaker. When I rely on interpersonal moral testimony, I can only be sure that the testimony is the outcome of whatever process it is you are following in moral deliberation. And so, I would be motivated by an explicit concern with moral rightness to take your testimony seriously. Alternatively, I could specify my request for testimony, asking you what it would take in this situation to strike the right balance, what really matters, etc. And, again, I suspect that many pessimists about moral testimony will be more inclined to think of such cases of reliance as less problematic than the ‘pure’ cases of reliance on moral testimony.

But wait—couldn’t this be the parity breaker we have been looking for all along? No, and for two reasons. First, it doesn’t track the difference between intra- and interpersonal moral testimony. While cases of interpersonal moral testimony might be more likely to involve an explicit concern for rightness, this need not necessarily be the case. Similarly, while many cases of intrapersonal moral testimony will not involve an explicit concern for rightness, but only a concern for rightness under another description, there is nothing about intrapersonal moral testimony that makes this so.

More importantly, however, it seems questionable that the content of the motivation should make a difference in the first place. For what exactly does a concern with the right-making features (i.e. the sort of motivation that is often described as ‘motivation by rightness *de re*’) come down to—if not an explicit concern with each of those features, i.e. a motivation by friendship, charitability, etc. *de dicto* (cf. Johnson King (2020b), 414)? And it is difficult to see why motivation by moral rightness *de dicto* should be objectionable while motivation by some right-making feature, say friendship, *de dicto* should not. Such an intermediate stopping point seems arbitrary, but going down ‘all the way’ in the metaphysical hierarchy seems overly demanding for praiseworthiness. Indeed, in the context of the debate around moral testimony, an intermediate stop would face further difficulties. For if such an intermediate stop at, say, fairness would be appropriate, then reliance on testimony about *fairness* would be permissible. But that is certainly not in the spirit of pessimism about moral testimony. The pessimism is supposed to be all-encompassing in the moral domain.

Considering all of this leaves the optimist about moral testimony in a much better dialectical position than they were in before. As we have seen, the optimist need not dig their heels in and argue that it is only an explicit concern about moral rightness that matters when it comes to moral motivation. Moral motivation can be praiseworthy in all kinds of ways—including an explicit concern about rightness, but also concerns regarding rightness under a different

description or indeed a concern about right-making features. Indeed, the optimist about moral testimony can wholeheartedly espouse the importance of a concern about the right-making features, for relying on someone else's moral testimony (or your past self) means taking steps to ensure that you will end up doing the right thing—it is just that the optimist also trusts other agents to track those right-making features even if they themselves are unable to double-check. The optimist about moral testimony properly construed, I suggest, is not primarily optimistic about moral testimony. Really, they are optimistic when it comes to other agent's moral capabilities.

3.4 Conclusion

In this chapter, I presented a novel argument for optimism about moral testimony. I suggested that reliance on your past self in making moral decisions is both commonplace and legitimate and argued that there is no relevant difference between such reliance and reliance on 'external' moral testimony. If so, then reliance on such interpersonal moral testimony is likewise legitimate.

Drawing on this result, I suggested some reasons for why reliance on your past self seems fine, particularly so from the perspective of pessimists about moral testimony. In relying on your past self, I argued, you are nonetheless showing adequate care about the things that pessimists about moral testimony usually value—right-making features of an action. Given the parity claim, I then tried to show that the very same thought in fact applies to interpersonal moral testimony as well, thus hoping to strike some common ground. This also provided a more nuanced understanding of optimism about moral testimony which becomes viable because there is more theoretical space available in the debate around moral motivation than the contemporary framing suggests.

Chapter 4

Moral Experts: Authorities *and* Advisors

Should non-experts consider moral experts' testimony as providing them with a preemptive reason for belief or as one reason amongst others? On the former proposal, the role of moral experts is to convey moral knowledge whereas their role is to aid in the development of moral understanding on the latter. In this chapter, I argue that the role of moral experts' testimony varies. Since neither proposal regarding the role of moral experts' testimony does justice to this fact, I propose a hybrid of the two proposals. By default, the aim of moral inquiry is moral understanding and non-experts should therefore treat moral experts as advisors which foster moral understanding. If, however, moral understanding is unavailable or only attainable at high risk, 'mere' moral knowledge becomes the aim of inquiry and non-experts should treat moral experts as authorities. Finally, I suggest that the aim of inquiry is in turn determined by the identities we take on in a domain. Among other things, taking on an identity just means being sufficiently involved in the domain to cultivate understanding. This explains why the default aim of moral inquiry is moral understanding—we simply cannot 'opt out' of our identity as moral agents.

4.1 Introduction

Suppose there are moral experts.¹ How should we—non-experts—take their moral testimony into account? For non-moral testimony, two dominant positions have evolved. Zagzebski (2012) argues that expert’s testimony provides non-experts with a preemptive reason for belief and a non-expert is thus well-advised to defer to the expert’s testimony. On this view, experts act as authorities. In contrast, Lackey (2018) argues that non-experts should consider expert’s testimony by taking it into account in their deliberation instead of merely deferring to the expert. On this view, experts act as advisors.

In this chapter, I first argue that the authority model presupposes knowledge as the (most important) aim of inquiry while the advisor model presupposes understanding as the (most important) aim of inquiry. The differences between the two models can therefore be traced back to different presuppositions regarding the aim of inquiry. Thus, it is the *aim* of inquiry that determines how non-experts should take expert’s testimony into account. Applying this proposal to moral experts, I argue that the aim of moral inquiry varies. Sometimes, the (most important) aim of moral inquiry is moral knowledge and, at other times, the (most important) aim of moral inquiry is moral understanding. Hence, in some contexts non-experts should treat moral experts as authorities and in others, they should treat moral experts as advisors.

I then suggest a framework for determining the aim of moral inquiry in a given situation. Roughly, I argue that we have an obligation qua being a moral agent to aim for moral understanding. However, this obligation is defeasible—if moral understanding is unattainable or comes only at high risk, moral knowledge becomes the most important aim of moral inquiry. The aim of moral inquiry, then, is determined by our identity as moral agents. Since we cannot opt out of this identity, understanding is the default aim of moral inquiry.²

We should thus, by default, treat moral experts as advisors because our default aim qua moral agents is moral understanding which is in turn best attained by following the advisor model. However, this default aim of moral inquiry is defeasible. There are times when we should treat moral experts as authorities. Thus, the answer to my initial question depends—there is no one size fits all solution.

¹It goes without saying that this is assumption is controversial. The contemporary debate on this subject goes back to Singer (1972), for a helpful overview (and distinctions of different kinds of moral expertise), cf. Driver (2013). For some problems with this debate as well as some discussion regarding the possibility of moral experts, cf. chapter 5.

²Naturally, the relevance of this result depends on the frequency the default aim is in fact defeated. I provide explicit arguments that reliance on moral testimony is permissible much more frequently than sometimes thought in chapter 2 and chapter 3.

4.2 Moral Experts

The question who counts as a moral expert is similarly controversial as the question about their existence. For the purposes of this chapter, it will thus be helpful to settle on a fairly minimal understanding of moral expertise. For this, I draw on Goldman (1999)'s account of expertise and apply the account to the moral domain. Goldman proposes a truth-linked account of experts (cf. Goldman (1999), 268 and Goldman (2001), 91–93). On this account, an expert's body of knowledge in a certain domain satisfies a comparative and a substantial requirement. According to the comparative requirement, the expert has more knowledge than most other agents in that domain. According to the substantial requirement, the expert surpasses a certain threshold of knowledge in that domain. I follow this proposal and summarise both requirements as the knowledge-condition. In addition, Goldman requires experts to be able to successfully deploy their knowledge with regards to new questions in the domain. Thus, one condition concerns the expert's knowledge whereas the other concerns the ability of the expert to deploy that knowledge.

For the purposes of this chapter, I assume that these conditions also apply to experts in the moral domain. As long as a metaethical theory is consistent with non-subjective truth-conditions for moral propositions, it is possible that an agent satisfies both conditions (cf. Cholbi (2007), 333). Hence, we need to assume the truth of some such metaethical theory for this understanding. In this chapter, I remain neutral towards the question whether there is one unified moral domain or whether there are several 'subdomains' within the moral domain. If you are more sympathetic to the former proposal, read 'moral domain', 'moral expert', 'moral knowledge', 'moral understanding', etc. as referring to a unified moral domain. If you are more sympathetic to the latter proposal, read these terms as referring to a specific moral subdomain. One result of my discussion of moral expertise is that their existence seems much more plausible if we think of morality as consisting of several 'subdomains'—it is easily conceivable that some of us have epistemic advantages when it comes to particular fields or topics (cf. my discussion in chapter 5 for some more direct arguments in favour of this diversified conception of moral expertise).

Following Goldman's account, a moral expert then has an extensive fund of knowledge in the moral domain as well as the ability to form new knowledge in the moral domain. While the knowledge-condition refers to the moral expert's moral knowledge, the referent of the ability-condition is less clear. I submit that this condition is plausibly captured by the expert's *understanding* of their moral knowledge. Recall Hills (2009)' explication of this notion in terms of abilities someone has with respect to a certain proposition:

If you understand why p (and q is why p), then in the right sort of circumstances, you can successfully:

- (i) follow an explanation of why p given by someone else;
- (ii) explain why p in your own words;

- (iii) draw the conclusion that p (or that probably p) from the information that q ;
- (iv) draw the conclusion that p' (or that probably p') from the information that q' (where p' and q' are similar to but not identical to p and q);
- (v) given the information that p , give the right explanation, q ;
- (vi) given the information that p' , give the right explanation, q' .

To understand why p , you have to have the abilities i–vi to at least some extent. (Hills (2009), 102–103)

The abilities Hills lists correspond to Goldman’s requirement to successfully deploy one’s knowledge with regards to new questions in the domain. Expanding Hills’s conception from a singular proposition to a body of knowledge, a moral expert is able to exhibit (i) to (vi) for all (or at least most) of their moral knowledge. An agent is a moral expert, then, if and only if they have an extensive fund of moral knowledge as well as understanding of (most of) their moral knowledge.

4.3 Authorities and Advisors

Even on the assumption that there are experts it is not obvious how non-experts should take their testimony into account. One option is to consider their testimony as providing non-experts with a preemptive reason, thus a non-expert is well-advised to simply defer to the expert’s testimony. On this view, experts are authorities. Another option is to take the expert’s testimony into account in one’s deliberation instead of accepting it preemptively. On this view, experts are advisors. These models are best understood as two extreme positions, leaving space for more moderate proposals (such as my own) in between. In the following, I show that the authority model is successful when stipulating knowledge as the (most important) aim of inquiry. In contrast, the advisor model is successful when stipulating understanding as the (most important) aim of inquiry. Therefore, which account to favour depends on the aim of inquiry.³

Let me begin with the account of experts as authorities, focusing on Zagzebski’s outline of such an account. Both for the formulation and the justification, Zagzebski draws on Raz (1988)’ account of political authority. She argues that we have a reason to treat some agents as epistemic authorities whose beliefs we should adopt preemptively:

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. (Zagzebski (2012), 107, my emphasis)

³Indeed, depending on which account we favour, there might also be a shift in who counts as a moral expert—as we will see, the rough direction is that the authority model would focus on the knowledge-condition in experts while the advisor model would focus on the ability-condition.

Two questions that are importantly interconnected arise. First, who are these authorities? Second, why should I treat their testimony in this way? Zagzebski's response to the second question sheds light on the first as well. She argues that I should treat an agent as an *epistemic* authority if that person is more likely than me to form a true belief:

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. (Zagzebski (2012), 110)

Consequently, I should treat those agents as epistemic authorities who are, regarding the domain in question, more reliable than me. For the purposes of this chapter, we need not worry about how much more reliable that agent has to be since we are interested in the question how non-experts should treat expert's testimony—by definition, the relationship of non-experts and experts is such that the latter are, because of their extensive fund of knowledge and their understanding thereof, sufficiently more reliable than the former.

Accordingly, on an understanding of experts as authorities, non-experts should treat expert's testimony as preemptive reason for belief. Zagzebski draws on the track record argument to make this point. Following this argument, treating the expert's testimony as a preemptive reason for belief gives non-experts the highest probability of reaching true beliefs (which are well justified by the expert's testimony). This argument, however, is sound only if we add the further assumption that *knowledge* is the (most important) aim of inquiry.⁴ Granting Zagzebski the validity of the track record argument, I focus on the assumption regarding the aim of inquiry in the following.

In this vein, Lackey worries about an expert who provides clearly false testimony (cf. Lackey (2018), 234–236). There are two versions of this worry. First, even if the non-expert identifies the testimony as clearly false, following the authority model, they should still accept the belief preemptively. Second, because of the preemptive treatment of the expert's belief, the non-expert might not even be able to recognise the testimony as clearly false. In both scenarios, the non-expert does not reason independently of the expert, because they treat the expert's testimony as preemptive. Lackey argues that this is not a coincidence: the authority model disincentives independent deliberations as these have no effect whatsoever if our final belief is determined by the expert's testimony. This problem also translates to the cases where a true belief is transmitted—even here, the belief will not be integrated into the reason-belief structure of the non-expert. Lackey's arguments thus amounts to the claim that the authority model at best provides the non-expert with *isolated* knowledge. However, she thinks that this focus on knowledge misses an important goal of inquiry:

⁴Kelp argues that knowledge is indeed the aim of inquiry, rejecting 'mere' true belief and justified beliefs as the alternatives (cf. Kelp (2021)). For our purposes, this discussion is not quite on point—really what we would be looking for is an argument in favour of 'mere' knowledge as opposed to, in particular, understanding.

For we might not be interested only in experts helping us reach the truth but in *how* they help us reach the truth. (Lackey (2018), 239, my emphasis)

The upshot of Lackey's worries regarding the authority model is that it provides the non-expert with a high degree of knowledge, but fails to provide them with understanding of that knowledge (cf. Hills's account above). Suppose p and p' are true and suppose further that someone who understands why p is true is able to tell that p' is true. If the non-expert learns that p via the expert's testimony and is then presented with a further testimony non- p' from the same expert, they recognise the falsity of non- p' only if they understand why p . However, since the authority model typically allows the non-expert to gain knowledge, and not understanding, they will be unable to reject the clearly false testimony non- p' . In short, the non-expert ends up accepting true beliefs without understanding why they are true.

Zagzebski briefly discusses this worry:

Would we think that the subject who gets to the truth through the direct use of her own powers deserves more credit for reaching the truth when that method is a less reliable strategy than taking the belief on authority? If you think the answer is yes, you must think that the value of believing by relying upon ourselves and not others is not just a constraint on the aim of true belief; it is an independent value. And it is not only independent, it *trumps* the value of true belief. (Zagzebski (2012), 117–118., my emphasis)

While Zagzebski dismisses the idea that there is a more important aim of inquiry than true belief, Lackey takes the idea seriously and uses it as the starting point for her own account of how non-experts should treat expert's testimony.

Based on her arguments against the authority model, Lackey proposes an advisor model according to which the expert's testimony is just one reason among others for the non-expert. On the advisor model, the non-expert is supposed to deliberate independently, taking the expert's testimony into account rather than simply accepting it preemptively. From the perspective of the expert:

I might explain the matter to you or offer arguments or other support for the belief at issue. I may be sensitive to your particular epistemic needs or concerns. I might be patient and open-minded in my treatment of you as an epistemic agent. In general, though, I am far more likely to *cultivate understanding* in you about the matter if I am your advisor rather than your authority. (Lackey (2018), 239, my emphasis)

On Lackey's account, expert's testimony has the same role as other reasons for believing in a given proposition. Thus, a non-expert should merely *consider* an expert's testimony in their deliberation rather than effectively replacing their deliberation with the testimony. The main reason to favour the advisor model over the authority model thus concerns the aim of inquiry. Lackey claims that *understanding* rather than knowledge is intrinsically valuable. The authority model might provide the non-expert with knowledge, but it fails to provide them with understanding.

In contrast, the advisor model is likely to provide the non-expert with understanding. Since understanding consists in abilities, the mere belief transfer is not sufficient for the non-expert to gain understanding. Rather, the non-expert is asked to evaluate arguments and reasons on their own and look for explanatory connections, thereby fostering understanding. On the advisor model, non-experts avoid absolute reliance on the authority without greatly diminishing their probability to get things right (though note, following the track record argument, they do diminish it at least slightly). Consider again the scenario in which p and p' are true and understanding p entails the ability to recognise the truth of p' . On the advisor model, the non-expert who is provided with the expert's testimony p not only knows but understands why p and is thus able to reject the clearly false testimony non- p' .⁵

However, consider scenarios in which understanding is simply not an aim of inquiry. You don't need to know why the showtime for the movie *Oppenheimer* in the *New Picture House* is 7.50pm. Even if there was some benefit to understanding it and an expert around to provide the relevant explanations, it seems like a waste of resources if all you need the time for is showing up at the theatre to watch the movie. Hence, there are scenarios in which the advisor model seems inapt, because its focus on understanding neglects the positive result of the authority model which ensures a high probability of knowledge.

Therefore, I submit, the question whether to adopt the authority or the advisor model is first and foremost a question regarding the most apt aim(s) of an inquiry. If we chose to follow the authority model, we favour the aim of gaining a high probability of knowledge. If we chose to follow the advisor model, we neglect this aim to some degree and decide to focus on understanding. Depending on our aim, one or the other model is preferable. It is this line of thought that I pursue in the remainder of this chapter.

4.4 The Hybrid Proposal

The worries with the authority model and the advisor model show that their difference does not concern the role of expert's testimony itself but rather the aim of inquiry. Zagzebski can admit that non-experts should treat experts as advisors if the (most important) aim of inquiry were understanding. We can then understand Zagzebski as holding that the latter assumption simply is, more often than not, false. Similarly, Lackey can admit that non-experts should treat experts as authorities if the (most important) aim of inquiry were knowledge. Again, we can understand Lackey as taking this latter assumption to be, more often than not, false. Hence, our interim conclusion from the previous section

⁵Of course, applying the advisor model requires a great deal of effort on behalf of the expert. Especially for the moral domain, we might think this entails that there will often be a great burden put on moral experts. Assuming that many moral experts will be those affected by injustices (as I suggest in chapter 5), applying the advisor model would require them to go through difficult issues with possibly ignorant inquirers. Cf. Berenstain (2016) for the underlying phenomenon of epistemic exploitation.

is that the role of an expert's testimony is determined by the aim of inquiry in the following way:

- (P1) If the (most important) aim of moral inquiry is moral knowledge, we should adopt the authority model for moral experts.
- (P2) If the (most important) aim of moral inquiry is moral understanding, we should adopt the advisor model for moral experts.

I take this list to be exhaustive, i.e. there are no other potential aims of moral inquiry that would justify a different model for moral experts.⁶ This is, in light of the understanding of moral experts I provided earlier in this chapter, plausible. Since moral experts satisfy both a knowledge-condition and an ability-condition (spelled out in terms of understanding), non-experts can plausibly acquire two things from moral experts. First, moral knowledge, second, moral understanding. On the authority model, non-experts (mainly) acquire the former, on the advisor model, non-experts (mainly) acquire the latter.

Sliwa (2012) argues that moral testimony sometimes is useful because it puts an end to one's own critical deliberations, and sometimes plays the role of any other reasons available to the agent (cf. Sliwa (2012), 183). Following this idea, we might interpret her as claiming that some occasions call for the authority model (where moral testimony potentially puts an end to own deliberations) and other occasions call for the advisor model (where moral testimony plays the role of any other reason). Given our interim conclusion, I propose that Sliwa is right precisely because the aim of moral inquiry varies. If the purpose of my asking for moral testimony is to put an end to my own deliberations, I am asking for moral knowledge. In contrast, if the purpose of my asking for moral testimony is to gain additional reasons for independent deliberation, I am aiming at moral understanding. Hence, sometimes the aim of moral inquiry is moral knowledge whereas it is moral understanding at other times:

- (P3) Sometimes, moral knowledge is the (most important) aim of moral inquiry and at other times, moral understanding is the (most important) aim of moral inquiry.

In conjunction with (P1) and (P2) we can conclude:

- (C1) Sometimes, we should adopt the authority model and sometimes we should adopt the advisor model.

⁶One might object that other relevant aims include the *practical* aims of inquiry, e.g. doing the right thing while the aims I am considering here are of a rather theoretical nature. However, they are tightly connected to practical aims. If you focus on gaining knowledge, then doing the right thing will become easily available. If you focus on gaining understanding instead, the practical aim connected is not just doing the right thing but doing it for the right reasons. Thus, there is an interesting parallel between theoretical aims of moral inquiry on the one side and practical aims on the other.

Let us call this the hybrid proposal. Depending on the aim of moral inquiry, non-experts sometimes should treat moral experts as authorities and sometimes non-experts should treat moral experts as advisors.⁷

4.4.1 The Aim of Moral Inquiry

In this section, I show that the hybrid proposal is plausible in its application in the moral domain. In particular, I demonstrate that the (most important) aim of moral inquiry sometimes is moral knowledge and moral understanding at other times. Further, I argue that one major factor influencing our intuition towards different cases of reliance on moral testimony is the respective aim of moral inquiry. Generally speaking, agents applying the authority model seem to be acting rightly only if the (most important) aim of ‘their’ moral inquiry is moral knowledge whereas agents applying the advisor model seem to be acting rightly only if the (most important) aim of ‘their’ moral inquiry is moral understanding. Agents who are mistaken about their aim of inquiry can thus go wrong in aiming at moral knowledge where in fact moral understanding would have been appropriate or in stubbornly aiming for moral understanding where the more modest aim of moral knowledge would have been appropriate instead.

I begin this discussion by providing two instances of moral testimony in which a non-expert follows the authority model. In the first, the agent seems to make a mistake by applying the model. In the second, they do not:

Claire has just been appointed as a judge and is very anxious to sentence people justly. [...] Luckily, she has a mentor, a more experienced judge, Judith, who has excellent judgment. Claire always consults with Judith and gives her decision in accordance with Judith’s guidelines, offering Judith’s explanation of why the sentence is just to the defendants. (Hills (2009), 110)

Susan’s friends have been playing pranks on a new girl in her class. Susan worries that they might be going too far and that they are bullying the girl. [...] Susan doesn’t know what to do. If her friends are being bullies, she should step in. But they are her friends and she doesn’t want to get them in trouble and she’s not quite sure whether what they are doing *is* bullying. Eventually, she decides to ask a friend for advice [and follows this advice without further deliberation]. (Sliwa (2012), 178, emphasis in original)

Hills claims that Claire is acting wrongly in following Judith’s judgment and Sliwa claims that Susan is doing the right thing in following the judgment of the friend she trusts. I have similar intuitions regarding both cases. Let us assume

⁷Friedman (Forthcoming) suggests a form of quietism about the aim of inquiry, holding that there is not much we can say about the aim of inquiry (cf. Friedman (Forthcoming)). I find it difficult to see, however, how this can be the full story. After all, most of her discussion suggests that there simply is no *one* answer to the question what the aim of inquiry is. Rather, I take it that many of her arguments ultimately support a form of pluralism, not unlike the one I suggest here.

that Claire and Susan are each right in their judgment that their mentor/friend is a moral expert on the matter of dispute whereas they themselves are not. Both Claire and Susan apply the authority model. I propose that we can explain why Claire is acting wrongly in doing so while Susan is not by drawing on the hybrid proposal. Claire's moral inquiry is such that its aim is moral understanding. Hence, she should apply the advisor rather than the authority model. In contrast, the aim of Susan's moral inquiry is moral knowledge wherefore she acts rightly in applying the authority model. Let me discuss both cases in detail.

Hills claims that Claire's dependence on Judith prohibits her from being a good judge. Crucially, she lacks moral understanding because she wholly relies on Judith's moral testimony. Hills suggests that Claire should instead exercise the capacity to base her beliefs about what is just on the actual reasons *why* it is just (cf. Hills (2009), 110). If this analysis is correct, then Claire's mistake is not in asking Judith, but in applying the authority model. It is the entailed preemptiveness of Judith's judgment for Claire's belief that is the problem because it prohibits Claire from gaining moral understanding. Instead, Claire should treat the moral testimony as one reason amongst others, thus exercising her own justice capacity. Why is the (most important) aim of this particular moral inquiry moral understanding? Arguably, judges should be experts in issues of justice (or come as close to that ideal as possible). If being an expert requires moral understanding and Claire strives to become one, she needs to acquire not just moral knowledge but moral understanding. Hence, the way Claire considers the moral expert's moral testimony is mistaken because the aim of her inquiry is moral understanding. Accordingly, she is acting wrongly in applying the authority model and should apply the advisor model instead.

Susan, on the other hand, seems to do the right thing in applying the authority model. Sliwa claims that Susan's reliance on her friend is justified since she cannot be certain about her own judgment due to possible biases towards her friends (cf. Sliwa (2012), 179). Susan seems to have exercised all the moral understanding available to her but is worried that her biases might render her judgment wrong. Since she cannot, for the moment, do away with her biases, she should, in order to increase her chances of getting things right, rely on her friend's moral testimony preemptively. In this particular case, the (most important) aim of moral inquiry is not moral understanding but moral knowledge. Hence, Susan is right in applying the authority model.

There are various other cases discussed in the literature to which similar remarks apply. The first class of cases, where the authority model is applied but should not be, involve agents who should strive for moral understanding. Unlike in the case of Judith above, this is mostly because the moral knowledge around which the case revolves is very basic and the protagonists, as mature moral agents, should be able to understand why a certain action is required rather than just knowing that it is. These cases involve a person wondering whether they should save a child even if they ruin their suit doing so (cf. Sliwa (2012), 176), a person who decides not to kill a person merely because of someone

else's moral testimony (cf. Sliwa (2012), 185), or a person wondering whether it is permissible to burn a cat for fun (cf. Groll and Decker (2014), 54).

On the other hand, there are various cases in which agents who apply the authority model seem to do the right thing. Two cases very parallel to the bullying case I discussed above concern a couple wondering about the permissibility of asking the wealthier parents for larger wedding contributions (Sliwa (2012), 177) or a journalist wondering whether a lie towards her family about a trip into a dangerous area is permissible (Sliwa (2012), 178). Again, these cases show that there are situations in which moral knowledge is the (most important) aim of moral inquiry, mainly because there are internal constraints on the attainability of moral understanding for the agents in question. Perhaps the most striking case concerns young children who accept norms simply because of their parents say-so (cf. Hills (2009), 98). Since young children are not (yet) mature moral agents, it is likely to be difficult or simply impossible for them to develop certain kinds of moral understanding. Hence, the aim of the respective moral inquiry cannot (yet) be moral understanding.

So far, I have been concerned with agents who follow the authority model in their moral inquiries. Let me now consider cases in which agents follow the advisor model. Again, I take it that there are cases in which the agents seem to act wrongly and others where they do not. Consider the following cases:

Sam, who is *white*, is walking through the city with their friend Jess, who is Black. Across the street, they observe a 'random' police check. Jess calls the incidence out for what it is—racial profiling—and asks Sam to intervene. Sam, however, does not intervene, instead asking Jess for further explanation, wanting to understand first why it is an incidence of racial profiling.

Sam, who is *white*, is walking through the city with their friend Jess, who is Black. As they talk, they come across the issue of racial profiling. While Sam doesn't understand the wrong of racial profiling at first, their conversation helps Sam to get there, allowing Sam to understand the wrong of racial profiling.

Sam seems to be doing something wrong in the first scenario, where the second seems more innocuous, even though they are applying the advisor model in both. What is wrong with applying the advisor model in the first scenario? Well—the time pressure. The intervention needs to happen *now*. The aim of moral inquiry is, because of the external time-constraint, not moral understanding, but moral knowledge. Conversely, without the time-constraint present (as in the second scenario), the aim of moral inquiry is moral understanding, not moral knowledge. As a mature moral agent, Sam should understand the wrongness of racial profiling. Note that the problem is not necessarily that Sam is entirely incapable of forming moral understanding. He might successfully gain understanding in both cases. However, he fails to gain moral understanding *in time* in the first case.⁸

⁸Unlike for the authority model, there is—to the best of my knowledge—no discussion of cases relying on something akin to the advisor model. This may be due to a tacit assumption

Based on all four cases, I want to end this discussion by drawing tentative conclusions regarding the question what determines the aim of moral inquiry which I support further by offering systematic reasons in the next section. Similar to Groll and Decker (2014), I take it that the moral domain is such that many issues require an agent to strive for moral understanding by default. The scenario involving the very basic moral truth that you should save a drowning child no matter the costs to your clothes illustrated this. Consequently, I suggest to conceive of moral understanding as the default aim of moral inquiry. Being a mature moral agent, you should, by default, strive for moral understanding.

However, there are reasons that make moral understanding in principle or at least in a given situation, unattainable. Consider the external constraint of time pressure or a high stakes decision. If you are in a situation in which you have to act fast and you have a moral expert by your side (as Sam does), preemptively believing the moral expert is the right thing to do because the (most important) aim of that particular moral inquiry is moral knowledge. Similarly, there might be agents who have internal constraints on their moral understanding. Possessing certain privileges may render the application of the advisor model very difficult because you struggle to give the proper weight to the experience of oppressed groups—as might be the case in one version of Sam’s case above. Likewise, biases may also preclude you from forming moral understanding. It may be that it is difficult or even impossible to form moral understanding on matters where you lack the experience of being on the receiving end of an injustice. All of these considerations provide you with reasons to abandon moral understanding as the (most important) aim of inquiry and opt for moral knowledge instead. Social location, then, may be an important internal factor constraining the aim of moral inquiry.

Thus, the following picture suggests itself. By default, moral understanding is the aim of moral inquiry and agents should apply the advisor model. However, if moral understanding is, for some reason or other, in principle or in a given situation unattainable or only attainable at high risk, moral knowledge is the aim of moral inquiry and agents should apply the authority model.⁹ In the next section, I argue that this is due to our non-optional identity as moral agents.

4.4.2 Refining the Hybrid Proposal

By default, agents should apply the advisor model towards moral experts because the default aim of moral inquiry is moral understanding. However, if there are serious constraints on the inquiry such that moral understanding cannot be the aim of inquiry, agents should apply the authority model instead because moral knowledge becomes the primary aim of that particular inquiry. So far, I

that gaining moral understanding is always innocuous. The cases discussed here suggest that such an assumption is misguided.

⁹Agents can, of course, be in error about the aim of a particular moral inquiry. There might thus be a procedural epistemic obligation to ensure that one is not in error before applying either model (cf. Rosen (2004), 301, for a discussion of the concept of procedural epistemic obligations).

have argued for this refined hybrid proposal only by an appeal to a few cases as well as the general structure of these cases. Still, you might wonder whether there is more to be said on *why* moral understanding would be the default aim of moral inquiry. In responding to this question, I also hope to shed light on the consequences for non-moral inquiry.

In their discussion of the difference between moral and non-moral testimony, Groll and Decker (2014) consider two conceptual reasons for such a difference—optionality and scope. According to optionality, morality is different to other domains in that there is no opting out. While it is possible to opt out of many of your identities, say that of your profession, it is not possible to opt out of morality. Declaring that you are opting out of morality does not free you of moral constraints (cf. Groll and Decker (2014), 62). The second reason, scope, refers again to the various identities we inhabit as agents. Here, the idea is that you can still be a good, say, physicist even if you rely on testimony of your colleagues a lot. The same, however, does not hold for morality. To wit, while some reliance on testimony may be problematic in other identities, reliance on testimony *always* seems problematic in morality (cf. Groll and Decker (2014), 63). Ultimately, Groll and Decker argue that neither of these reasons supports a principled difference between moral and non-moral testimony. While I agree with this diagnosis, I think investigating both reasons more closely still allows us to show that there is *some* difference between moral and non-moral inquiry, albeit less principled than has been suggested in the literature (a reading which Groll and Decker seem to reject). In particular, because of optionality and scope, understanding is the default aim for all agents in the moral domain while it is the default aim in other domains only relative to the identities an agent chooses to take on.

Let us turn to optionality first. While Groll and Decker agree that one cannot opt out of morality, they argue that the same holds for other domains such as rationality or perception. Essentially, this list extends to all the identities humans have qua being human. Since opting out is not unique to the moral domain, they conclude that there is no principled difference between moral and non-moral testimony. Rather, optionality would suggest a principled difference between the identities we *chose* to take on and the identities we *have* to take on qua being human. While this insight is sufficient for their argument against the asymmetry between moral and non-moral testimony, it does provide further insights for the present discussion. After all, in claiming that moral understanding is the default aim of inquiry in the moral domain, I do not need to claim that this does not hold for other domains as well.

Following Groll and Decker, understanding seems to be the reasonable aim if we have (for some reason or other) taken up an identity that requires us to make judgments and decisions about the domain in question. Among other things, taking on an identity just means being sufficiently involved in the domain to cultivate understanding. While there are some identities that we have to take on as humans, there may be others we can choose to take on—I mentioned the example of our profession above, but we can think more loosely of identities we

might take on, the identity of a loving partner, a caring parent, etc.¹⁰ Concerning all these domains, it seems reasonable to think that we fulfil this identity only if our primary aim in these domains is not ‘mere’ knowledge, but understanding. You would hardly count as a loving partner if you made decisions on your relationship purely based on the testimony of others, not going in any way to understand the underlying issues. Of course, even in domains where you take up an identity, you should sometimes put the aim of understanding to the side—precisely under the conditions outlined in the previous section.¹¹

While there is no principled difference between morality and other domains, optionality still has effects on the default aim of inquiry. So far, I argued that taking on an identity creates an obligation to strive for understanding in the respective domain. Thus, whenever we take on such an identity, understanding becomes the default aim in the domain—be it morality or any other. In the following, drawing on Korsgaard (1996)’s account of the sources of normativity, I suggest that our moral identity, however, is one we cannot opt out of. Therefore, understanding is the default aim in the moral domain for everyone.¹²

Korsgaard (1996) argues that we cannot opt out of our moral identity by basing moral obligations on the reflective structure of human consciousness. This reflective structure enables us to make laws for ourselves and simultaneously gives rise to practical identities. Only once we occupy such practical identities can we act for reasons (cf. Korsgaard (1996), 100–113). While most of these practical identities are contingent (being a loving partner, being a caring parent, etc.), Korsgaard argues that we are necessarily governed by *some* conception of our practical identity as to not be governed by such a conception is to lose your grip on yourself. This necessity, however, does not arise due to some particular practical identity, but originates in your humanity itself—precisely from the fact that your consciousness is structured reflectively. In this sense then, all of our particular identities depend on our identity as human and thus on the value of humanity (cf. Korsgaard (1996), 113–128).

¹⁰Actually, it is not exactly clear what it means to take up a certain identity—how actively do we need to be invested in a domain to count as taking up an identity in the respective domain? Fortunately, we do not need to settle that question for our purposes here, as we ‘only’ need the assumption that all of us take up an identity in the moral domain. Note that this need not even mean to actively endorse the identity you are taking up. Cf. Korsgaard (1996) for an account of practical identities where the moral identity in particular is one that needs no active endorsement.

¹¹If you put the aim of understanding aside for a particular question, say the judgment Claire makes following Judith’s guidance, can you later pick the same question up and strive to understand the underlying reasons? Of course—indeed, it seems like you have some obligation to do so, resulting from the default aim of moral understanding. So, the notion of ‘inquiry’ I rely on here is very fine-grained and goal-oriented. When it comes to the judgment Claire has to make, the most important aim is moral knowledge and thus the authority model is apt. This does not, however, preclude Claire from taking the inquiry up later again and aiming for moral understanding. Callahan (2018) worries that it might nonetheless disincentive aiming for moral understanding later. Cf. chapter 2 for a discussion of her worry (cf. Callahan (2018), 454).

¹²Korsgaard’s account is particularly helpful here because it allows for the strongest formulation of non-optionality. However, even if one were to reject her particular account, I take it that non-optionality still enjoys a lot of intuitive pull.

The final step in Korsgaard’s argument is to show that valuing your own humanity commits you to valuing the humanity of others as well. For this, she draws on the earlier insight that our identities, including our identity as human, give rise to reasons for actions. She then argues that the space of linguistic consciousness, in which reasons exist, is inherently public. Therefore, reflective consciousness about ourselves is not of a special kind and varies only in degree to reflective consciousness about others. Thus, insofar as we value humanity in ourselves, we are committed to valuing it in others, as well (cf. Korsgaard (1996), 132–145).

On Korsgaard’s account, then, it is the reflective structure of human consciousness that creates moral obligations. If we are to occupy any practical identity, we also have to occupy the practical identity of being human, valuing humanity in ourselves and others—this *moral* identity becomes “inescapable” (Korsgaard (1996), 130).¹³ And so, because we cannot opt out of morality, the default aim is moral understanding for all of us. This allows for a generalisation not possible in opt-out domains such as one’s profession, one’s emphasis on a relationship, etc. In these domains, too, understanding is the default aim—but it is the default aim only if an agent chooses to take on the respective identity.

This gives us a solid justification for the first part of the refined hybrid proposal—for all of us, moral understanding is the primary aim of of moral inquiry. However, this does not yet give us a reason to think that the second part of the refined hybrid proposal holds. According to this second part, it is admissible to go for moral knowledge if moral understanding is unavailable.

This is where Groll and Decker’s discussion around scope will prove helpful (they use the term ‘role’ to refer what I discuss under the heading ‘identity’):

The idea, recall, is just that when it comes to non-moral roles, it is often perfectly possible for someone to fulfill the role—and even fulfill it well—while still accepting a great deal on the basis of testimony. But, the objection goes, the same is not true for the moral realm, where reliance on testimony always militates against one’s adequately fulfilling the role of moral agent. (Groll and Decker (2014), 68–69)

Perhaps unsurprisingly, I agree with Groll and Decker that scope does not hold in the absolute sense outlined above. It is not the case that reliance on testimony (i.e. being satisfied with pure knowledge) precludes one from fulfilling the

¹³There is one exception to this inescapability: if humans decided that human life was worthless, i.e. that no practical identity was worth taking up, Korsgaard thinks that no practical identity and thus no moral identity arises (cf. Korsgaard (1996), 254). A similar move does *not* work for what we might think of as ‘immoral’ identities. Korsgaard considers the example of a Mafioso who has a code of strength and honour, a loss of which would threaten his identity as a Mafioso. This identity is immoral because it is in regular conflict with any moral obligations the Mafioso might have. The reason that the Mafioso is still committed to the more fundamental moral identity is that one of the rules of reflection is “that we should never stop reflecting until we have reached a satisfactory answer, one that admits of no further questioning. [...F]ollowing that rule would have led the Mafioso to morality, and, since he was reflecting, he ought to have followed it, and therefore he ought to have arrived there. His obligation to be a good person is therefore *deeper* than his obligation to stick to his code” (Korsgaard (1996), 258).

identity of a moral agent. To argue for this, Groll and Decker point to cases of moral dilemmas, difficult evaluations of how various moral considerations compete, etc. While they admit that there may be something suboptimal about an agent unable to figure out what the right response is in such cases, the agent is certainly not violating their obligations as a moral agent if they seek out testimony from other agents (cf. Groll and Decker (2014), 71–72). I think we can even add that the agent in question may be fulfilling their identity as a moral agent perfectly well precisely because they recognised their own limitations and decided to seek out testimony for this very reason.

This line of thought should be familiar from the previous section where I argued that internal and external constraints may cause an agent to rationally change their aim of moral inquiry to the more modest aim of moral knowledge. Internal constraints include biases or a lack of certain experiences while external constraints include time constraints or high stakes.

Thus, while there is a correct insight in the underlying idea behind scope, it simply does not go all the way. While we may, as moral agents, often be required to aim for moral understanding to fulfil our identity as moral agents properly, this obligation is by no means indefeasible. Recall the formulation of the refined hybrid proposal—if moral understanding is unattainable or only attainable at high risk, then aiming for moral knowledge (e.g. via testimony) is sufficient. In fact, moral knowledge becomes the primary aim of moral inquiry in such situations so that aiming for moral understanding instead may even violate the obligations we have qua moral agents.

The aim of this section was to provide a rationale for the refined hybrid proposal. I have argued that this rationale consists primarily in the identity as moral agents we take on qua moral agent. This non-optionality of morality accounts for the default aim of moral understanding. However, this default aim is defeasible in situations where moral understanding is unattainable or only attainable at high risk. In such cases, moral knowledge becomes the primary aim of moral inquiry instead.

Hopefully, the discussion of optionality and scope also sheds light on non-moral inquiry. For these, the default aim of inquiry depends on whether we choose to take on an identity in the respective domain. If we do, then the default (but defeasible) aim of inquiry is understanding. If we do not, the primary aim of inquiry is just knowledge with understanding as a bonus not required of us. This allows us to apply the findings of this chapter to non-moral inquiry. As for moral inquiry, neither the authority nor the advisor model captures all there is—again, they are best applied in tandem in the form of a hybrid proposal. The aim of inquiry determines which model is to be applied, and the identities we take on in turn determine the aim of inquiry.

4.5 Conclusion

Throughout this chapter, I discussed the role of moral experts' testimony for non-experts. I illustrated two proposals and argued that neither is, on its own,

sufficient to properly capture the role of moral experts' testimony. First, I discussed the authority model according to which moral experts' testimony provides a preemptive reason for belief. I argued that this model presupposes moral knowledge as the aim of moral inquiry. Second, I discussed the advisor model according to which moral experts' testimony should play the role of any other reason for belief the non-expert has. I argued that this model presupposes moral understanding as the aim of moral inquiry.

Drawing on the insight that the role of moral experts' testimony depends on the aim of moral inquiry, I proposed a hybrid of the two models. Following this proposal, non-experts should apply the authority model if moral knowledge is the aim of their moral inquiry and the advisor model if moral understanding is the aim of their moral inquiry. I demonstrated the plausibility of this hybrid proposal by discussing examples of moral inquiries in which agents succeeded or failed in employing the model appropriate to the aim of moral inquiry. Drawing on these cases, I suggested that the aim of moral inquiry can be determined as follows: moral understanding is the default aim of any moral inquiry while moral knowledge becomes the aim of a moral inquiry if moral understanding is either unattainable or only attainable at high risk.

Finally, I provided a rationale for this refined version of the hybrid proposal. I argued that our non-optional identity as moral agents entails moral understanding as the default aim of moral inquiry and suggested that this entailment is defeasible in some cases where the agent is either internally or externally constrained.

Chapter 5

Moral Expertise and Moral Philosophy

In this chapter, I consider and reject Cholbi (2018)'s proposal for identifying moral experts. He suggests that theoretical knowledge—knowledge of the correct moral theory—is a central element of moral expertise and could thus be considered as a starting point for an independent check in morality. In response, I draw on Star (2015)'s work around moral worth who argues that we can reliably and non-accidentally act for the right reasons even if we are ignorant of the correct moral theory. I apply this discussion to Cholbi's proposal and argue that, similar to the case of moral worth, moral expertise comes in many ways—only one of which is via theoretical knowledge. I argue that Cholbi's suggestion is symptomatic for the traditional conception of moral expertise—a conception that focuses too much on theory and is concerned too little with practice.

5.1 Introduction

On my first walk through St Andrews with my supervisor, we came across the Pends, a pair of fourteenth-century arches at the east end of South Street. Local legend has it that the archways will collapse if a true genius walks through:

At the east of South Street, there's a pair of 14th Century arches known as The Pends, which were part of a gateway to the walled enclosure surrounding the cathedral. [...] I discover there is a superstition that if a 'true genius' walks through the Pends, the archways will collapse—Just as well I don't attempt it. (Ritchie (2015), 30)

The superstition had such an impact that no other than John Stuart Mill, then rector of the university, refused to walk through the Pends fearing their potential collapse (cf. Kunkler (2015), 30). If the superstition were true, walking through these archways would provide an excellent *independent check* for true genius. Without relying on the qualities that constitute true genius, you would be able to tell whether someone is a true genius simply by observing The Pends.

But we need not venture into the realm of legends to find independent checks. Consider an astronomer, who, relying on their astronomical knowledge, 'predicts' a solar eclipse for some particular day. Without any astronomical knowledge of our own, we can independently verify their expertise by waiting for the day in question and observing the sky (cf. Goldman (2001), 106).

Notably, however, there is no such independent check in morality:

It is harder to see how one might calibrate the accuracy or reliability of one's *moral* judgment; one lacks the relevant kind of *independent* access to the moral facts. [...]t seems as though one could do so only by engaging in first-order moral reasoning and deliberation of one's own. It is thus unsurprising that clear and unequivocal evidence that someone else's moral judgment is significantly more reliable than one's own is hard to come by. (McGrath (2009), 334, my emphasis)

McGrath makes two related observations here. First, we cannot access moral facts other than by moral theorising. Hence, there is no independent check in morality. Second, due to the lack of an independent check, it is impossible to figure out whether someone is significantly more reliable than ourselves, or, in other words, whether someone is a moral expert.

Cholbi (2018) disagrees. He suggests that theoretical knowledge—knowledge of the correct moral theory—is a central element of moral expertise and could thus be considered as a starting point for an independent check in morality (cf. Cholbi (2018), 85). I don't think this is correct. But I think it is incorrect in an interesting way: in this chapter, I argue that Cholbi's suggestion is symptomatic of the traditional conception of moral expertise—a conception that focuses too much on theory and is concerned too little with practice. My hope is that this provides some positive aspects of a new conception of moral expertise—one that may be easier to check independently. In particular, I suggest that the question 'Who is a moral expert?' might ultimately be the wrong question to ask. Rather,

we should wonder “Who is an expert on this particular moral matter?” where moral matters comprise fields less broad than ‘the moral domain’ but still broad enough to be interesting enough for systematic analysis—e.g. racism, friendship, etc. And here, it is much less compelling to think that it is impossible to identify experts in—and these, I suggest, just are moral experts, properly understood. Identifying experts in such subdomains might even be possible using quasi-independent checks, by paying close attention to the particular standpoint an agent occupies.

I start by discussing why an independent check for moral expertise is desirable not just qua contributing to the debate around moral expertise, but also for each individual’s moral reasoning. Next, I introduce Cholbi (2018)’s proposal in detail, paying close attention to the asymmetry between moral and non-moral domains he introduces. Afterwards, I draw on Star (2015)’s work around moral worth who argues that we can act for the right reasons even if we are ignorant of the correct moral theory. I apply this discussion to Cholbi’s proposal and argue that, similar to the case of moral worth, moral expertise comes in many ways—only one of which is via theoretical knowledge. I then suggest that thinking of moral expertise in this way provides us with a much more optimistic outlook towards our ability to identify such experts.

5.2 Why an Independent Check?

Recall McGrath’s observation about the lack of independent check for morality—she argues that it is not possible to gain clear evidence for other people’s reliability in moral matters. Setting intellectual curiosity aside, you might think that this is not particularly worrisome. After all, if there is no independent check and hence no way to assess reliability in moral matters, we could instead simply suspend judgment on the reliability of other agents when it comes to morality.

There are at least two reasons why this is unattractive. First, the required suspension of judgment would be more far-reaching than just on our credibility judgment of *other* agents. As Sliwa (2012) observes, if we lack an independent way of assessing the reliability in moral matters of other agents, then there is also no reason to suppose that we can evaluate *our own* reliability in moral matters:

[...T]he lack of an ‘independent check’ doesn’t just preclude [an agent] from assessing other people’s reliability. It also makes it impossible to assess her own reliability—she is no more in a position to check whether she herself got it right than she is in a position to check whether someone else got it right. (Sliwa (2012), 191)

Sliwa’s observation shows how radically the lack of an independent check undermines our ability to assess reliability when it comes to moral matters.

The second reason to look for a genuine independent check for moral expertise comes from a more practical perspective. Even if we are not epistemically

justified in making judgments about the reliability of agents in moral matters, most of us nonetheless do make such judgments—it is just that they lack a solid foundation. What is worse, Dular (2017) argues that our judgments of moral reliability are not just unfounded, but likely to be systematically distorted—in favour of agents in privileged positions.

Dular traces this back to two interconnected reasons, both resulting from the oppressiveness of our society. First, under conditions of oppression, members of subordinate groups will more often feel uncertain in moral matters than members of dominant groups and are thus more likely to look to someone else for moral advice. Second, under conditions of oppression, members of dominant groups will be ascribed more competence than is justified. Hence, our confidence and competence judgments are systematically distorted.

Dular’s argument for the first claim—that members of subordinate groups are more likely to be uncertain in moral matters than members of dominant groups—stems from research on stereotype threats. Stereotype threat describes a phenomenon where an individual feels the risk of confirming a negative stereotype about their own social group (cf. Steele and Aronson (1995)).¹ The most striking consequence of stereotype threat is that it undermines one’s performance in domains where negative stereotypes about one’s social group exist. Next, Dular draws on a study by Schmader et al. (2004) that stereotype threat also undermines one’s *confidence* even if it does not undermine one’s *performance* in the domain. So much for general argument as to why someone experiencing stereotype threat may have too little confidence about their performance in some domain. Next, Dular (2017) argues that members of subordinate groups experience more negative stereotypes than members of dominant groups. Hence, she concludes, members of subordinate groups are likely to have less confidence in their judgments than members of dominant groups.

The next step in her argument is to show that there are negative *moral* stereotypes about members of subordinate groups. If so, she suggests, then it follows that members of subordinate groups have less confidence than members of dominant groups when it comes to moral matters, thus giving rise to Dular’s first criticism. To show this, she draws on the traditional conception of morality as a supposedly rational and impartial endeavour and argues that it conflicts with stereotypes about women who are stereotyped as “emotional, biased, vain, weak and incapable of making tough or serious decisions” (Dular (2017), 217). Next, she considers people of colour in the US as another subordinate group. Here, she draws on research showing that crimes committed by people of colour are more likely to be attributed to negative attitudinal and personality traits than crimes committed by their *white* counterparts (cf. Bridges and Steen (1998) for the respective study). At the very least then, Dular provides an argument that negative moral stereotypes exist along the lines of race and gender.

Her arguments for the systematically distorted ascription of competence follow a similar line. Again, she first draws on more general research demonstrating

¹Though it is well-documented phenomenon, some worry that the relevance of stereotype threat may nonetheless be overrated due to a publication bias (cf. e.g. Flore and Wicherts (2015)).

that people are more likely to attribute the successes of a person to luck if they are a member of a subordinate group than if they are a member of a dominant group (cf. Biernat and Kobrynowicz (1997)). Dular combines the finding of this study with the study by Bridges and Steen (1998) cited above to conclude that similar results are to be expected for the moral domain. Since people are more likely to attribute a bad moral character to members of subordinate groups than to members of dominant groups, it is also much more likely that moral mistakes are attributed to the bad moral character for members of subordinate groups and to unfortunate circumstances for members of dominant groups. Dular expects these findings to carry over to judgments in a similar manner. Hence, our beliefs about the competence of agents in the moral domains are likely to be systematically distorted, favouring members of dominant groups.

A note on Dular's discussion is in place. While Dular is concerned only with *moral* testimony, her argument extends to cases of non-moral testimony as well. In fact, when outlining why we tend to underestimate the reliability of members of subordinate groups, she draws on research on the estimation of competence *in general* and then goes on to show why this general finding *also* holds for morality. If so, then there is nothing especially problematic about *moral* testimony (pace the general assumption in the literature on moral testimony), at least in this respect. Still, one might wonder whether Dular's focus on moral testimony can nonetheless be justified. I think it can—drawing on the lack of an independent check in the moral domain (cf. McGrath (2009), 334). Thus, while a non-distorted way of estimating someone else's reliability in non-moral matters may at least in principle be available (though it may be that it is employed only rarely), the same does not hold for the moral domain. This provides at least one way in which the special treatment of morality is justified. If we cannot independently verify our estimates of someone's reliability when it comes to moral matters and we know that our estimates are systematically distorted, then it seems that we should best give up on moral testimony as a resource.

Nonetheless, this observation supports my suspicion that our response to Dular's argument need not entail that we should disregard moral testimony entirely. Instead, we could try to find better practices for deciding whom we can trust when it comes to moral testimony. If independent checks allow us to circumvent our own biases in non-moral domains, we might wonder whether a similar tool is available in the moral domain.

5.3 Theoretical Knowledge

Cholbi (2018) argues that theoretical knowledge—knowledge of the correct moral theory—is a prerequisite or at least an important element of moral expertise, and might thus get us closer to an independent check in the moral domain. He claims that moral expertise consists at least in part of theoretical knowledge and holds that this knowledge is essential to the expert's social function of giving moral advice (cf. Cholbi (2018), 71).

To argue for this position, Cholbi distinguishes two conditions a moral expert has to satisfy. First, an *epistemic condition* according to which their first-order moral knowledge must significantly exceed that of their non-expert counterparts (this suggestion is in line with the early writing on expertise, e.g. Goldman (1999), 268 and Goldman (2001), 91–92). Second, a *testimonial condition* according to which it is this high level of first-order moral knowledge that justifies their treatment as reliable moral advisors (cf. Cholbi (2018), 76–77). It is this second condition that is crucial to Cholbi’s argument.

Cholbi justifies the relevance of the testimonial condition in the moral domain as follows. In non-moral inquiry, especially when we are concerned with an expert’s testimony, the most important factor in considering whom to trust is predictive success. If a theory is successful in predicting observations, this provides a good reason to accept it. “The justificatory relationship”, Cholbi notes, “runs from the observation to theory” (Cholbi (2018), 79):

Suppose that some scientific theory T implies some observation O, and that O is observed in some experimental setting. O thereby confirms T. (Cholbi (2018), 79)

Thus, mere reliability is enough in most non-moral domains, there need not be a proper grounding in theory for an observation to be accepted. This model, however, Cholbi argues, does not hold for moral inquiry. In fact, for moral inquiry, the justificatory relationship is reversed—it runs from theory to observation. The truth of some first-order moral theory predicts particular intuitions all the same, but more importantly it *justifies* the corresponding judgments:

As in the scientific case, moral theories are ‘tested’ against first-order judgments. But the fact that we accept a moral theory because it implies plausible first-order moral claims does not entail that the justificatory relationship runs from the first-order claims to the theory. Rather, the truth of the theory justifies the first-order claims. Moral inquiry aims at something more than predictive success, namely, explanatory grounding of our first-moral claims. (Cholbi (2018), 80)

From this, Cholbi concludes that we cannot detach the reasons for our first-order moral judgments from the moral theory that implies them. Therefore, Cholbi argues, a moral expert can contribute to moral inquiry qua moral expert only insofar as they possess knowledge of the correct moral theory. Otherwise, the testimonial condition is undercut (cf. Cholbi (2018), 84).

Cholbi takes care to place the relevance of theoretical knowledge in moral experts into context. Taking it to be a requirement for moral expertise, he thinks that it is critical for the identification of a moral expert:

More importantly, an expert’s inability to give any theoretical accounting of her testimony regarding first-order moral questions would, I propose, be a ‘red flag’ alerting non-experts that the expert’s claim to reliable first-order moral knowledge should be second guessed. (Cholbi (2018), 85)

I want to take this remark as the starting point for my discussion around the relevance of theoretical knowledge for moral expertise ultimately aiming to establish that theoretical knowledge cannot be a necessary condition to check someone's moral expertise.

5.4 Derivative and Fundamental Moral Reasons

Emphasising the importance of the proper grounding of one's moral judgments and actions is not a new idea. It has been most influential when it comes to discussions around the *moral worth* of an action. Moral worth requires not just performing the right action, but performing it *for the right reasons*. In order to engage with Cholbi's argument, it is therefore instructive to make a detour to the debate around moral worth—after all, both this debate and Cholbi's proposal centre around grounding one's moral judgments in the right way. The question I want to press for now is whether moral worth requires knowledge of the fundamental right-making features for an action (i.e. knowledge of the correct first-order moral theory as suggested as the requirement for moral expertise by Cholbi).

The position that knowledge of the fundamental right-making features is required certainly has some initial attraction. The underlying idea is that performing the right action for the right reasons requires knowledge of these reasons. A natural understanding of these reasons is to take them to be the right-making features of an action, i.e. the most fundamental features that render the act in question right. If, however, knowledge of these fundamental reasons is required for performing an action for the right reasons, then so is knowledge of the correct moral theory—after all, you will only be in a position to recognise the fundamental moral reasons if you are equipped with the correct moral theory, or so the argument goes.

Star (2015) argues that knowledge of the correct first-order moral theory cannot be required for moral worth. For if it were required, it would invalidate too many of our ordinary judgments about the moral worth of actions of people who do not know or care about the correct first-order moral theory.² Star suggests that the acting for the right reasons clause is satisfied also by acting for the right *derivative* reasons:

Suppose, just for the sake of simplicity, that the correct ethical theory is hedonistic utilitarianism. This theory would have it that the only fundamental reasons are facts about pain and pleasure. [...] At the level of basic theory, hedonistic utilitarianism says nothing at all about promises. Yet, very plausibly, given contingent facts about the world, the fact that

²Though with a slightly different motivation, a similar idea can be found in Sidgwick (1908). Sidgwick argues that knowing the correct ethical theory—utilitarianism, in his view—might be detrimental to pursuing it. Instead, pursuing secondary goods to happiness, such as virtue, truth, freedom or beauty might be more conducive to the primary good, happiness (cf. Sidgwick (1908), 405–406).

one has promised to do something is (quite strong) evidence, in a situation where one is provided with an opportunity to fulfill a promise, that one ought to fulfill the promise. What ultimately makes the fact that one has promised evidence that one ought to fulfill one's promise is that it is generally true that fulfilling a promise leads to an increase in pleasure or a decrease in pain. (Star (2015), 15–16)

Imagine a person fulfilling a promise simply because they care about promises and believe them to be of moral value. Following Star, their action would have moral worth (as seems intuitively true) even though it fails to respond to the fundamental right-making features of their actions, which, by hypothesis, is the truth of hedonistic utilitarianism. Even though the agent in question is ignorant of this fundamental right-making feature, their action has moral worth via their responsiveness to the *derivative* right-making features.

Some clarification on the notion of derivative reasons is in place, for it is not enough that responding to a derivative reason produces reliable results. Arpaly (2002) asks us to imagine a world in which an invisible hand makes the motive of profit a reliable indicator of the right action (similar to Star's conception of promises above). Still, Arpaly maintains, we would not consider actions performed on the basis of a motive for profit as having moral worth (Arpaly (2002), 225). What distinguishes the cases?

Johnson King (2020a) offers one plausible explanation. Moral worth requires not just acting for the right reasons but *deliberately* acting for the right reasons (Johnson King (2020a), 201–202). If so, then we can provide two versions of Arpaly's thought experiment. First, an act performed purely for the motive of profit thus lacking moral worth. Second, an act performed for the motive of profit *because it is believed to play a significant moral role* (in the sense of a derivative reason—it is a reliable indicator of the right action). The morally worthy action is thus not just one performed in response to the right reasons (whether they are fundamental or derivative) but in response to the right reasons recognised *as* right reasons.

Note that this discussion gets real traction only when considering derivative reasons. Acting on the basis of the right reasons without recognising them as such in the case of *fundamental* reasons, i.e. the actual right- or wrongmaking features, is rare.³ For derivative reasons, however, things can come apart quite easily—because they are not fundamental, they can often be the right reasons to respond to without at the same time being recognised as such—consider Arpaly's example above. Knowledge of the right fundamental reasons (and acting upon them) is thus sufficient for performing actions of moral worth because such knowledge usually comes with recognising them as moral reasons. However, actions of moral worth can also be performed when acting 'merely' on the basis of the right derivative reasons—it is just that we need a clarification here to

³It is not impossible—we could, perhaps, receive a financial incentive for acting on the basis of a fundamental reason, thus acting for the right fundamental reason but doing so without necessarily recognising it as the right fundamental reason. These cases, however, are at least non-standard.

exclude Arpaly's cases: we need to respond to the right derivative reasons while at the same time *recognising* them as moral reasons.

The main upshot of Star's discussion for our purposes here is that knowledge of the correct first-order moral theory is not necessary for performing actions of moral worth. Of course, this claim does not yet contradict Cholbi's argument that knowledge of the correct first-order moral theory is necessary for fulfilling the social role of a moral expert. Arguably, many people will satisfy the conditions for moral worth while fewer people will satisfy the conditions for moral expertise. Still, recognising the value derivative reasons can have allows us to question the validity of Cholbi's claim regarding moral expertise.

5.5 Derivative Reasons and Moral Expertise

The overarching question is whether moral expertise requires knowledge of the correct first-order moral theory, i.e. knowledge of the fundamental right-making features of an action. Cholbi answers this question to the positive, arguing that only knowledge of the fundamental reasons enables non-experts to justifiably recognise someone as a moral expert—being a moral expert is not sufficient to be justifiably recognised as such.

The first problem this strategy encounters is that non-experts typically do not know the correct first-order moral theory—hence the need for expert advice. Therefore, it is not clear how they could identify those agents who have knowledge of the correct first-order moral theory (recall McGrath's initial worry for this point). Cholbi's solution to this problem is to weaken the conditions for the identification of moral experts such that the important identifier becomes whether or not a putative moral expert can provide some justification via fundamental reasons to back up their moral advice. If they cannot, we are justified in questioning their claim to moral expertise thus rejecting them at least in their social role as a moral expert.

This strategy, however, comes with an undue elitisation of moral expertise. We have learned from Star's discussion that the group of people doing inquiry into the fundamental moral reasons is typically restricted to moral philosophers. Star's way of avoiding an undue elitisation of moral inquiry is to point out that responding to fundamental moral reasons is but one way of responding to the correct moral reasons—with responding to derivative moral reasons as an equally valid way of responding to the correct moral reasons. In contrast, Cholbi's account straightforwardly excludes this option from the beginning—after all, *only* fundamental reasons can play the role required for Cholbi's account.

Cholbi's response to this issue is to highlight the positive aspects of this understanding of moral expertise. He emphasises that even the average moral agent is fairly reliable in moral matters. Thus, following his epistemic condition, moral experts have to have some relative advantage to this (already high) level of moral knowledge:

In my estimation, that moral controversies or dilemmas receive so much popular and scholarly attention should not obscure that first-order moral knowledge, particularly with respect to relatively straightforward moral phenomena, is pretty widely distributed among human moral agents. [...] Moral experts must therefore have *significantly* more first-order moral knowledge than this. (Cholbi (2018), 76, my emphasis).

If Cholbi is correct in his optimistic evaluation of moral agents, then the high requirements for moral expertise may seem justified. However, especially when granting that the average moral agent is fairly reliable, I take it that this can only be because there are *multiple* ways to excel in moral inquiry and thus multiple ways of satisfying the epistemic condition. Following Cholbi, one of these ways may be knowledge of the fundamental moral reasons. However, other ways include knowledge of (or responsiveness to) derivative reasons.

This, I take it, is where Cholbi's argument goes wrong. He rightly argues that predictive power is insufficient for the identification of experts in the moral domain. From this, however, he moves on to the claim that therefore knowledge of fundamental reasons is required as an identifier of moral expertise. With this leap, he overlooks a middle ground between mere predictive power and all-out knowledge of fundamental moral reasons—knowledge of (or responsiveness to) derivative reasons.

If there are multiple ways of excelling in moral inquiry, only one of which is via knowledge of fundamental reasons, then requiring knowledge of the correct first-order moral theory systematically excludes a substantial portion of moral experts while it over-emphasises the importance (and the findings) of another sub-group within moral experts—namely moral philosophers.

5.6 Traditional Conceptions of Moral Expertise

So far, I was mostly concerned with a particular proposal for an independent check for moral expertise. However, I think my rejection of theoretical knowledge as a requirement for moral expertise is best understood not as a response to a singular proposal, but as a rejection of a widely shared assumption on moral expertise—namely the assumption that it is *moral philosophers* who have a special claim to moral expertise:

Moral philosophers have, then, certain advantages which could make them, relative to those who lack these advantages, experts in matters of morals. Of course, to be moral experts, it would be necessary for moral philosophers to do some fact-finding on whatever issue they were considering. (Singer (1972), 117)

On the picture Singer is painting, moral philosophy is mostly concerned with developing abstract theories about right and wrong. Following the development of such a theory, we only need to do “some fact-finding” to find out how the theory applies in real world cases. The advantage moral philosophers are claiming according to Singer, then, has morality as a whole as its scope.

I think that this picture is severely mistaken. Morality is too vast a domain for any one standpoint to sensibly claim an epistemic advantage over as a whole. This mistaken assumption had the effect that the quest for moral experts has, at least in the philosophical literature, too often been focused on ourselves as moral philosophers. When wondering whether there are moral experts, we often just asked whether *we* are moral experts (cf. e.g. Archard (2011) and Gordon (2014) for some particularly telling titles, ‘Why moral philosophers are not and should not be moral experts’ and ‘Moral Philosophers Are Moral Experts!’). Now, I assume that nobody would explicitly uphold the equivocation of these two questions, but sure enough—there has been an implicit understanding that the two questions are at least intimately connected. I think that this equivocation is unfortunate, but explainable.

After all, when looking for moral experts, we have been looking for agents occupying a standpoint with a particular epistemic advantage in moral matters *overall*. The subject matter over which moral experts purportedly claim an advantage is just this: the moral domain. And indeed, the prompt ‘Who is a moral expert?’ naturally lends itself to this interpretation. I suggest that we do better when refining the scope, not considering morality as a whole but rather its various subfields and questions. This also lends itself to the natural occurrences of requests for moral expertise. Rare are the situations in which an agent is interested in morality overall and accordingly motivated to seek out this general expertise. More often, we are interested in a specific moral question and motivated to seek out guidance concerning this particular question or subfield of morality. Starting from this point, it is less appealing to think that moral philosophers would ultimately be the expert relative to some subfield.⁴

Note that even if moral philosophers *could* claim an advantage relative to others when it comes to the entirety of the moral domain, this is far from rendering them the expert for some subfield. For suppose it were true that when considering the moral domain, moral philosophers have an advantage over others, e.g. because of their theoretical knowledge. This assumption is compatible with it being the case that for each subfield of morality, there are individuals who are better placed than moral philosophers to answer the question of relevance in the field (hence the need for “some fact-finding”). Of course, on the assumption that moral philosophers are right in claiming an overall advantage when it comes to the moral domain, these individuals could not be the same in each field—but this result should be far from troubling. After all, for each moral question, it is likely that there is someone who is better suited than the moral philosophers to respond to the issue at hand.

To be sure, I am not claiming that moral philosophers have no kind of expertise in moral matters. It is just that their expertise lies in the width of morality, in their theoretical knowledge, not necessarily some subfield within

⁴My claim here does not require the moral domain to be disuniform in any deep sense. Nonetheless, it is certainly compatible with such an approach where e.g. the uniformity of the semantics and/or ontology of the moral domain or the uniformity of first-order moral theory is questioned (cf. e.g. Sinnott-Armstrong and Wheatley (2012) and Mesel (2019) for the former suggestion, Brännmark (2016) for the latter).

morality. In these subfields, however, fundamental reasons are of less epistemic significance than knowledge of or responsiveness to derivative reasons—these do the crucial work of grounding someone’s specific knowledge in the subfield. And when appealing to moral testimony, we are usually interested in some very specific subfield. In such cases, the perspective of moral philosophers will likely not be epistemically advantageous relative to an expert specialised in the field.

Who, under this new understanding of the question, is a moral expert, then? I am afraid that giving up the unity of the question as I suggest means that there is no simple answer available any longer. After all, the question ‘Who is a moral expert?’ now receives a different response for any particular subject matter. I think this is not a drawback of the account proposed—rather, it captures the unease many have expressed about the existence of moral experts (there is no unified answer available) while simultaneously allowing for the existence of agents who have, on a particular moral matter, a privileged position to others.

One promising way of finding an answer to the newly uncovered questions is thus to consider the various standpoints we take up on different moral matters. Since moral matters in all usual cases affect agents, we arrive at a putative class of experts—those affected by the matter at hand. For example, Kolers explicitly suggests that those who are least well off with respect to a political issue are best placed to give advice (cf. Kolers (2016), 103). This thought is not novel, but reflected in the long tradition of standpoint epistemology. Roughly, in considering a matter affecting a group, we should consider *their* perspective on the matter because it is informed by an epistemic privilege about the matter. Such epistemic privilege is gained not merely by being affected by an issue, but engaging with the issue in an appropriate way—that is, standpoints are not bestowed automatically but acquired (cf. Anderson (2020), section 2, for an outline of standpoint theory and its history).⁵

As is suggested on the traditional picture of moral expertise, it takes work to become a moral expert. However, the work is not confined to that of (academic) moral philosophy. Rather, it is engagement with the injustices of our world that renders one to take up the position of a moral expert—albeit ‘only’ with respect to the issues one engages in. But this does not limit the approach, rather it does justice to the complex injustices of the world we live in.

5.7 Conclusion

I have considered a proposal for identifying moral experts that focused on theoretical knowledge—knowledge of the correct moral theory. I raised some worries for this proposal, in particular drawing on the fact that while it may describe one way to excel in moral inquiry, it fails to take other ways of excelling in moral inquiry into account. This finding is particularly troubling because the proposal takes into account only *our* way (as moral philosophers) of doing moral inquiry.

⁵However, being affected by the issue is at least usually considered to be a necessary condition for acquiring a standpoint. Cf. Tilton (Forthcoming) for an argument against this necessity claim.

This, I argued, is symptomatic of the debate around moral expertise, though perhaps not unsurprising. The question ‘Who is a moral expert?’ is most naturally read as asking for an individual (or a class of individuals) who excel in all aspects of moral inquiry. But this is far from the only way of doing well in moral inquiry—especially once we consider that requests for moral testimony are always specific and thus specific to some subfield within the moral domain. Here, it is more likely possible to excel in moral inquiry also without theoretical knowledge. Responsiveness to derivative reasons is all that is needed.

I thus suggest to understand the question ‘Who is a moral expert?’ to be concerned with experts in relevant subfields of morality, e.g. moral matters such as racism, friendship, etc. And here, or so I hope is intuitive, it is much more compelling to think that there are experts in the field. We were just looking in the wrong place.

Chapter 6

Moral Testimony and Blame

Suppose I rely on someone else's moral testimony and end up doing the wrong thing. What, if anything, can we say about the allocation of blame in such cases of reliance on moral testimony? For pessimists about moral testimony, the answer seems straightforward: since I should not have relied on the testimony in the first place, I am the main suspect when it comes to allocating blame. However, unlike in non-testimony cases, there are at least two other options: blaming the testifier, and blaming both the testifier and the advisee. In this chapter, I consider these options and argue that both the wrongdoing agent and the testifier can be blameworthy—and can, in particular, legitimately be blamed by the wronged agent. I draw attention to the fact that taking both agents seriously as actors entails conceiving of both as *responsible* agents—in their deferring and in their testifying. This helps us to circumvent worries about shirking responsibility when it comes to reliance on moral testimony: relying on someone else's advice does not get you off the hook. At the same time, it also has some consequences for testifiers: sharing your moral beliefs, I suggest, can mean taking some responsibility for the actions taken on the basis of those beliefs—even if their enactment is outside your control. Relying on and sharing moral testimony, then, is as serious a part of moral discourse and practice as any other form of deliberation. I suggest that pessimism about moral testimony cannot do justice to this result while optimism about moral testimony can.

6.1 Introduction

I don't like surprise parties. I really don't, there is nothing at all I find appealing about them. Now, many of my friends know this about me and yet, every now and then, someone will throw a surprise party for me. I take it that this is morally objectionable. Yet, or so it seems, it isn't such a clear-cut issue as I think it is: many people, including some of my friends, are evidently convinced that it is permissible to throw unwanted surprise parties.

Suppose you are one of these people and despite knowing of my aversion to surprise parties, you go ahead and throw one for me. If I am right and it is indeed wrong to do so, then you wronged me and can be blameworthy for doing so. What is more, it seems that I, in particular, can appropriately blame you for throwing the surprise party.

So far, so good. But now imagine the situation were somewhat different. Again, you know that I dislike surprise parties, but you are uncertain whether this means that you cannot throw a surprise party for me (you *really* enjoy throwing surprise parties). Being uncertain about the matter, you turn to a friend of yours and ask them for advice. They tell you that it is in fact permissible to throw even unwanted surprise parties. Relying on their advice, you decide to throw a surprise party for me.

Again, if I am right about the permissibility of throwing unwanted surprise parties, I have been wronged. But by whom? It does not seem appealing to think that only you are to blame—after all, you only acted on your friend's advice. Neither does it seem appealing to locate all of the blame in your friend—it was *you*, after all, who decided to seek out their advice and follow through on it. Perhaps, then, my blame should be 'distributed' among the two of you. But if so, how? In blaming *both* you and your friend, am I now blaming 'more' than before? Are there different ways in which I blame the two of you—perhaps blaming you in a practical way, but your friend in an epistemic way? If so, were these two dimensions of blame present also when you came to the judgment on your own?

Hopefully, these initial considerations illustrate that there is an open question about our practices of blaming (and, relatedly, praising) when it comes to relying on other's moral advice. And considering this question, or so I hope, is not just interesting in its own right, but also sheds light on blame and praise in situations unrelated to moral testimony. Until recently, such a discussion would have had difficulty to get off the ground. One commonly shared assumption was that reliance on moral testimony is illegitimate anyway. Moral testimony, as Sliwa put it about ten years ago, "has been getting a bad name in the recent literature" (Sliwa (2012), 175). And if reliance on moral testimony is illegitimate, why bother discussing allocations of blame and praise when an agent relies on moral testimony? After all, the agent shouldn't have sought out moral testimony in the first place.

However, things seem to have changed over the last decade: there is a growing number of critics of such pessimism about moral testimony (cf. e.g. Sliwa

(2012) herself, Groll and Decker (2014), Lillehammer (2014), Wiland (2017), McShane (2018), as well as a recent book, Wiland (2021)). I myself am sympathetic to many of the arguments put forth by these so-called optimists about moral testimony and hope to have added to them in the previous chapters. But of course, the matter is by no means settled: just as there have been criticisms of pessimism about moral testimony, there have been extensive defences thereof, and I fear that the optimists are still in the minority (for just some recent defences of pessimism about moral testimony, cf. e.g. Hills (2009), Crisp (2014), Howell (2014), Fletcher (2016), Mogensen (2017) and Callahan (2018)).

Nonetheless, the recent surge in arguments against pessimism about moral testimony allows us to question whether reliance on moral testimony is in itself morally objectionable and thus allows us to ask questions we could not ask before—or at least, questions we could not sensibly ask before. In this chapter, it is the question of blame in cases of reliance on moral testimony that I pursue. I do so along two axes of investigation. The first is to consider intuitive responses to the allocation of blame in cases of moral testimony. Throughout this chapter, I suggest that these intuitions vary by context, not allowing for a one-size-fits-all kind of answer. The second axis is the question of how well optimism and pessimism can or cannot capture this analysis. This chapter is structured along the possible loci of blame—starting with the advisee and continuing with the testifier themselves. I argue that both agents can sometimes be held responsible. Optimism about moral testimony, I suggest, can accommodate this result while pessimism about moral testimony is committed to the sort of one-size-fits-all answer that seems implausible considering the phenomenon under investigation.

6.2 Blaming You

While I noted above that it might not seem appropriate to blame *only* you for the unwanted surprise party, it does seem that you are at least somewhat blameworthy for your action. Let us assume that you satisfy the standard criteria of moral agency and that you were free to act as you did. You were not, in particular, coerced into acting as you did by your testifier. So, you fulfil the general criteria for being held responsible.

For the pessimist about moral testimony, there is a straightforward explanation for the blame you receive: not only did you bring about a wrong, you did so basing your action on an illegitimate reasoning process. You violated, as it were, the hallmark of pessimism about moral testimony: you should not rely on other agent's moral testimony precisely because it fails to properly ground your moral beliefs/actions. For the pessimist, then, it is easy to explain why at least some of the blame (and, as we will see in the next section, perhaps all of the blame) should be directed towards you, the deferring agent.

However, things seem different for the optimist about moral testimony. After all, or so some optimists about moral testimony might initially be inclined to argue, seeking out moral testimony shows that you *do* care adequately enough not to be blamed. After all, wasn't the whole point of defending reliance on

moral testimony that such reliance is not an indicator of a defective moral character? Insofar as the optimistic project about moral testimony is to render reliance on moral testimony permissible, it might seem that holding the agent relying on moral testimony responsible is counterproductive.

I think this is too quick, but it helps to clarify the optimist's position about moral testimony. A worry pessimists about moral testimony have often had with reliance on moral testimony is that the agent in question is trying to shirk responsibility for the resulting action (cf. Sliwa (2012), 188–189 for an outline of the worry as well as a response in the spirit of what follows below). If we hold that moral testimony always exculpates in this way, then the worry about shirking responsibility seems to be on point: after all, we would essentially grant the premise of the worry. I take it this shows that the optimist about moral testimony *cannot* want moral testimony to be exculpating on its own because that would preclude a sensible response to the shirking responsibility objection.

On the other hand, I think we also should not be led to the other extreme, namely that reliance on moral testimony never exculpates, as e.g. the pessimist would have it—after all, if the optimist is right, then moral testimony can provide us with moral knowledge and it is, at least in principle, permissible to rely on moral testimony. A position in this vicinity is articulated by Harman (2019), who argues that reliance on moral testimony is in general permissible, but does not exculpate when going wrong.

Harman is, I think it is fair to say, an optimist about moral testimony. So why does she hold that moral testimony, when gone wrong, can *never* exculpate? Part of it is her idiosyncratic view of blame, as we will see below, but I think there is also an interesting relation to some of the motivation put forth in favour of pessimism about moral testimony and so it is worth engaging with the proposal in detail. Harman believes, in general, that false moral views cannot exculpate. Even though “ethics is hard” (the title of a 2014 draft paper of hers), she argues that ethical evidence is in principle available to all moral agents and so having a false moral belief must be a result of inadequate care, thus leading to the result that false moral beliefs cannot exculpate. Coupled with her belief that moral testimony *can* provide moral knowledge, however, it seems as if a tension arises. After all, if moral testimony can provide justified true moral beliefs, surely it can also provide justified *false* moral beliefs?

No, argues Harman, and points to an asymmetry about the ways in which testimony provides justification in the moral domain. Roughly, because Harman thinks that ordinary moral agents have a lot of evidence for true moral claims at their disposal, testimony *can* provide a justification for true moral beliefs while it *cannot* provide a justification for false moral beliefs since the justification would be undermined by the evidence at the agent's disposal (cf. Harman (2019), 178–180).

To get rid of the tension, one might just reject Harman's idiosyncratic view that false moral beliefs can never exculpate. Interestingly, however, the motivation for her view is surprisingly similar to some of the motivations pessimists about moral testimony have at their disposal. Reliance on moral testimony, or so they hold, demonstrates a failure with respect to an important moral

achievement, moral understanding (cf. Nickel (2001), Hopkins (2007) and Hills (2009)), a defect in the deferring agent's moral character (cf. Howell (2014) and Mogensen (2017)) or a psychological peculiarity (cf. Fletcher (2016) and Callahan (2018)). All of these motivations are importantly different, yet it seems that they all share a common assumption similar to Harman's view about the moral domain: in principle, you *could* figure things out and so you *should*—and if you cannot, well that's on you (and we're going to blame you for it).

I don't find either of these two extreme responses—moral testimony *always* exculpates (a caricature of the optimist about moral testimony) and moral testimony *never* exculpates (pessimism about moral testimony, but also Harman (2019))—compelling. The more moderate suggestion I would like to make on behalf of the optimist about moral testimony is this: reliance on moral testimony *can* exculpate, but it doesn't have to. To take some of the considerations from earlier into account: seeking out moral testimony *in order* to shirk responsibility is problematic. Our criticism of the 'shirking responsibly' objection should be that, in general, you can be the target of blame when relying on moral testimony because you do hold responsibility—seeking out moral testimony is an exercise of moral agency. It is precisely *because* we optimists think relying on testimony is not giving up agency that we should think that it is possible to be blamed when relying on moral testimony. The difference to the pessimist about reliance on moral testimony is rather that the deferring agent is not automatically blameworthy because seeking out moral testimony *can* be exemplify showing adequate care.

A helpful comparison here will be the situation where you throw the unwanted surprise party not based on someone else's moral advice, but following your own (independent) deliberation about the matter. The difference to that situation is your way of coming to the moral belief grounding your action. Usually, when assessing someone's blameworthiness for an action, we are interested in how conscientious they were in forming their underlying belief. If you are careless in forming your moral judgment we tend to think that blame is more appropriate than when you tried your best to get things right.

Usually, then, you are blameworthy for your inadequate care—and you can still be blameworthy for *that* if you seek out moral testimony. It is just that you are not blameworthy *qua* seeking out moral testimony because seeking out moral testimony is not in itself (pace the pessimist) a sign of inadequate care. If you seek out moral testimony precisely because you care about getting things right, then you certainly do not exhibit a lack of adequate care. If, however, you seek out moral testimony because you are lazy and do not care enough about the matter, then you are still subject to blame, but that is because of your inadequate care exemplified in your seeking out moral testimony, not in seeking out moral testimony itself. In fact, I take it that this is the kind of situation many pessimists about moral testimony have in mind when pronouncing reliance on moral testimony to be problematic.

In the present case, let us assume that you reached out to your friend because you thought the issue at hand is difficult and you did *not* want to make a careless decision about it. If, in addition, you have good reasons to assume that your

testifier is reliable, it does seem that you sincerely tried to get things right. In ‘usual’ cases (i.e. those where you do not rely on moral testimony), assessing the level to which you conscientiously came to your moral judgment would be enough to evaluate this variable for the level of blame that is appropriate. In cases involving moral testimony, however, this does not seem to be the full story. After all, we now have a new potential ‘target’ for my blame to evaluate: your testifier.

Before getting to that, we should note two ways in which you are nonetheless inextricably linked to the action (and the corresponding blame). First, in the perhaps obvious way that you performed the action. Without you, there would have been no surprise party and thus no wrongdoing. By acting (even if it is acting on the basis of someone else’s moral testimony), the act is at least attributable to you and you are thus potentially blameworthy if the act ‘goes wrong’. This much is parallel to cases where you do not rely on moral testimony.

Second, you took the initiative to seek out the testimony and if the above discussion holds, then it seems plausible to think that you have an obligation to make sure that the testimony you receive is in fact reliable. To satisfy this condition, you might even have some obligation to think about the matter yourself.¹ Pessimists about moral testimony might interject exactly at this point and declare that this very finding shows why reliance on moral testimony is impermissible. This, however, precludes another possibility of securing the testimony’s reliability, and that is by making sure that your testifier is in fact reliable on the matter and it should at least be a theoretical possibility to do so in other ways than checking for yourself—else we are begging the question (cf. McGrath (2009) for some worries regarding the feasibility of such checks and my discussion in chapter 5 for some ways out).

Let us assume that none of these undermining factors are at play here. Your seeking out testimony was done responsibly and you tried your best to get things right. All that is left for me to blame you for, then, is your being an ‘agent of wrongdoing’—in that you brought about a (preventable) wrong. If this is correct, then it does seem indeed that you are no more or less blameworthy than you are in the situation where you try to figure things out for yourself (assuming a similar level of conscientious judgment). But, again, there seems to be more to say in cases of moral testimony, and that is to say more about your testifier.

6.3 Blaming Your Testifier

There is definitely a sense in which your testifier is ‘complicit’ in the wrong you did to me. After all, were it not for them, you might have avoided your

¹At least in some cases, there might be a ‘should have known’-norm in place here. For some pieces of moral advice, it seems that we might reasonably ask of a moral agent that they come to know them on their own. However, this does not seem to track the difference between moral and non-moral testimony, but rather between what Groll and Decker identify as the difference between ‘normal’ and ‘non-normal’ knowledge (cf. Groll and Decker (2014)).

wrongdoing altogether. The question is whether your testifier might also be *morally* responsible, i.e. here blameworthy, for the act. Furthermore, as I indicated above, we should distinguish such potential blameworthiness from the question whether *I* can appropriately blame your testifier. As Smith argues, “the question whether an agent is responsible and culpable for an action or attitude, and the question whether [I] should actively blame her in some way for it, are different and to a certain extent independent questions” (Smith (2007), 472–273).²

Let us begin with this second question before returning to the first. *If* the testifier, too, can be subject to blame, can *I* blame them? Finding an answer requires us to think through the question whether I have the standing to blame your testifier. For such standing to obtain, there is usually a requirement of ‘being affected’ by the other person’s actions:

While we may not need any special authority to morally disapprove of another person on the basis of her objectionable attitudes or behavior, it seems that specific sorts of anger, resentment, and disappointment are open only to those *directly* wronged by the person in question. (Smith (2007), 479, my emphasis)

Even though I was not wronged directly *by* your testifier, I am nonetheless affected by their advice. In other words, it seems that your testifier’s advice becomes ‘my business’ at the latest when its enactment (by you) wrongs me. Interestingly, this might be the case even if there was no explicit intention on behalf of the testifier to make their moral belief my business.

In fact, we can even suppose that the friend you turn to for their advice is entirely unrelated to me. When asking for their guidance, you might have presented the matter in an entirely general way. Of course, your friend is nonetheless aware that you might act on their advice (though even that is outside their control), but they have no idea that *I* would be on the receiving end of such actions.

In this case, it might seem that there is something off about the idea of me blaming your testifier. While providing testimony may come with the awareness that the testimony-receiving agent might act on the basis of that testimony, the testifier has no control over such actions. From the perspective of the testifier, there is an interesting parallel to cases of moral luck where some think that factors outside of an agent’s control can nonetheless make a difference in our judgment of their responsibility (cf. Nelkin (2023)). Likewise, the testifier might here be held responsible for something that is, to a significant extent, outside of their control. Suppose two testifiers give out advice and of the two testimony-receiving agents only one acts on that advice. Are the testifiers to be held

²And it is the latter question that is of interest to here. If you think that testifying alone cannot be sufficient for being morally responsible (i.e. in the sense of Smith’s first question), the answer to the second question will be to the negative as well. So, the question I am mainly interested in here is this: *if* testifying alone (or testifying in conjunction with someone acting on your testimony) is sufficient for being morally responsible, does that secure the appropriateness of blame on behalf of the wronged person?

responsible to the same extent? Arguably, your stance on the issue of moral luck is likely to commit you to a similar position here. If you think that our testifying agents are to be held responsible only for testifying itself (no matter whether someone actually puts that testimony into practice), you are likely to ground that claim in holding that you are not to be held responsible for factors outside of your control. By contrast, if you think that the degree to which our testifying agents can be blameworthy differs, you seem to accept that also factors outside of our control can make a difference for our moral assessment.

There is a way in which this link becomes relevant only insofar as we consider the question whether *I* can appropriately blame your testifier. After all, the additional blame one of the testifiers might receive is now arguably coming from me. Upholding that I can appropriately blame the testifier whose wrong advice led to the wrongdoing seems to commit us to the position that giving moral testimony is, in this sense, like moral luck. If, however, you think that I gain no special standing to blame upon ending up on the receiving end of the wrongdoing, it is likely that you think that cases of moral luck can be explained away, eventually arguing that agents can only be held responsible for factors within their control.

A final way of approaching this question is to consider *what* the testifier would be blamed for, and this brings us back to the first question—are they to be blamed in the first place? Are they supposed to be blamed simply for holding a moral belief? I must admit that I have difficulty of grasping what this would even mean—although not across the board. Basu (2019) argues that there are at least some moral beliefs that you can be blameworthy for just in virtue of holding them. Most prominently, she discusses racist beliefs. Holding such beliefs, Basu argues, is not just wrong because they might result from flawed ways of reasoning (‘upstream’ problems) or because they result in morally wrong actions (‘downstream’ problems), but wrong because of the *content* of such beliefs (cf. Basu (2019), 2509–2512). While this seems plausible for the beliefs she considers, it seems plausible precisely because they are immoral beliefs *about* a particular individual in virtue of belonging to a social group or about the social group itself—but the same argument does not seem to translate moral beliefs in general and our case of moral advice in particular.

If they are not to be blamed for holding a moral belief, perhaps they are to be blamed for *sharing* the belief?

6.3.1 Passing the Buck

In cases of testimony, it seems that you, as the testimony-receiving agent, have a special standing to blame your testifier, for they testified not just to anyone, but to *you*, in particular (though it is questionable whether you yourself are subject to some wrongdoing). You trusted your testifier in their judgment and their mistaken judgment led to your wrongdoing. If I start blaming you for throwing the surprise party, it seems that one intuitive response for you is to ‘pass the buck’—it seems reasonable for you to cite the moral advice you received, perhaps in conjunction with the reasons you had to trust your testifier on the matter.

This is not a novel suggestion, but is rather familiar from the literature on testimony more generally. On ‘epistemic buck-passing’, Goldberg (2011) writes:

Suppose that hearer H accepts speaker S’s testimony that p, under conditions in which H had the epistemic right to accept that testimony; that some individual T later queries H regarding the truth of H’s testimony-based belief that p; that, in response, H exhausts all of her reasons for regarding S’s testimony as trustworthy; and that even so T remains unsatisfied. In this situation H is epistemically entitled—is within her epistemic rights—to pass the epistemic buck on to S (by representing S as having more in the way of epistemic support for the truth of p). Suppose [...] it turns out that S’s testimony to this effect had insufficient epistemic support. In this situation H is entitled—is within her epistemic rights—to *blame* S for the insufficient epistemic support of her (H’s) own belief. (Goldberg (2011), 178, my emphasis)

Of course, Goldberg is concerned with purely *epistemic* buck-passing, and there might be differences to the present case. Note, however, that the analysis should nonetheless translate to the epistemic dimension of our case. Following this analysis, there is no ‘need’ for *me* to blame your testifier—all I should do is direct my blame towards you and you can pass the buck onto your testifier, or, alternatively, blame them independently of my blaming you.

Both of these options would circumvent some of the issues under discussion essentially because they conceive of testimony-giving as an action in itself—an action which is subject to our usual blame- and praiseworthiness conditions. In virtue of this, they would solve some of the issues brought up so far. The case of blaming an agent who acted on the basis of another’s moral testimony would simply not be a special case after all.

Both ways of detailing the proposal, however, also bring some problems of their own along. Let us start with the idea of literally ‘passing on’ the blame though I should note that I am not even sure what would happen in the agent ‘passing on’ the blame psychologically. If there is something intuitive about the idea of passing on the blame, shouldn’t your reaction to being blamed by me be to explain your reliance on your testifiers advice and point out that the blame is thus *not* appropriately allocated to you, but should be directed towards your testifier (for that is exactly what you intend to do later)? But if so, why think that *you* should blame your testifier as opposed to me, the eventual target of your wrongdoing?³

I think this way of detailing the proposal is not appropriate for our present case, and the worry I have relates to the problem just outlined. At least in some cases, the blame the testifier receives can be greater than the blame you receive.

³This brings up an interesting point. Really, there seem to be three things we might blame any agent for: their wrong action, their way of coming to the underlying judgment *and* the underlying judgment. In the case of testimony, these two latter parts are distinct: *you* chose your testifier, but *your testifier* ‘provided’ you with the judgment. In usual cases of blame, these latter two parts will go hand in hand. I take it that these two parts correspond to what Goldberg (2006) distinguishes as the justifications you have to accept the testimony and the total epistemic support you have for the claim (cf. Goldberg (2006), 137–138).

These will be cases where your seeking out moral advice was done responsibly whereas (unbeknownst to you) your testifier showed inadequate care in coming to their belief. But if so, it makes little sense to think that what is going on is that you are merely passing on my blame, for where is the additional blame supposed to come from in these cases?

There is, however, another way of filling out the details of the story, and I think that while it takes the concept of ‘passing the epistemic buck’ less literally, it still coheres well with Goldberg’s suggestion. That is, we can understand your blaming your testifier as being in principle independent of my blaming you—thus accounting for cases in which your testifier receives more blame than you. Note that both the stronger and weaker interpretation of this buck passing nonetheless provide a convenient solution for the issue of standing to blame outlined in the previous section. While it might be unclear whether *I* have the standing to blame your testifier for testifying, you certainly do.

There are some complications for the independence-reading though: if your blaming your testifier is independent of my blaming you, it seems that you would have had the same standing to blame them even if their testimony did not result in your wrongdoing (e.g. if you didn’t act on their testimony but still received and trusted it). After all, the blame would be *for testifying*. But there seems to be a difference between merely receiving the testimony and putting the testimony into action—if you put their judgment into action, you become liable for the consequences in a way that you do not when you do not act on their judgment. Distinguishing blameworthiness and standing to blame might provide a solution here: perhaps your testifier is blameworthy for testifying all along, but you only get the standing to blame them for testifying once you put their testimony into action.

Another complication is that considering things from your testifier’s perspective, their reaction to you blaming them and me blaming them is likely to be different. For suppose you confront them about their mistaken testimony, detailing how you trusted them and how that led to your wrongdoing, wouldn’t one intuitive response from your testifier be “well, you didn’t have to listen to me”? And this seems very much in line with the considerations in favour of blaming you as opposed to your testifier in the earlier section—*you* are the one who acted, and you did not do so on their behalf, but by your own volition.

In contrast, the same sort of shrugging off does not work if *I* confronted your testifier with the wrong done to me. Their analogous response “well, they didn’t have to listen to me” lacks the force it had previously—because, whether your testifier likes it or not, you *did* listen (and I had no choice in whether they did or not), and so their testimony is involved in the wrong brought upon me.

It is for this reason that I don’t think that the blame your testifier receives can be entirely independent of me (though I am sympathetic to an approach where at least some of the blame—namely the epistemic dimension as outlined—is indeed quite independent of me). At the same time, the blame is not merely for testifying, but indeed for testifying *to you*, who in turn wronged me. To see why I have the standing to blame *only your* testifier (and not others holding the same (false) belief), consider the following case. Suppose two agents have a

mistaken moral belief, formed on the basis of inadequate care. Both agents, for example, might believe that throwing unwanted surprise parties is permissible. Now, if you turn for advice only to one of the two agents, it seems, following the discussion in this section, that I can legitimately blame them. The same, however, does not go for the other testifier whose testimony never affected me. There are, therefore, some restrictions on blame in place: namely, at least an indirect involvement in the wrong brought about.

6.3.2 Optimism and Pessimism

Whatever your stance on the above considerations, it seems to me that only accepting optimism about moral testimony allows for a fruitful debate about these complications. After all, by the optimists light, it is in principle legitimate to draw and rely on someone else's moral testimony, thus opening the door for blame being allocated both on the advisee and the testifier while also allowing that at least sometimes, both agents can be exculpated. For the pessimist, however, things seem to look different.

Pessimism about moral testimony, recall, is the idea that *reliance* on moral testimony is illegitimate. However, it is unclear what follows for *giving out* moral testimony. Note first that the illegitimacy of reliance on moral testimony does not automatically entail the illegitimacy of giving out moral testimony. For consider an analogy. Suppose you organise an art gallery, displaying various works of art from local artists. Mindful of their rights, you decide to prohibit visitors from taking photographs of the artworks. For the visitors, then, it becomes illegitimate to take photographs of the artworks. But that same rule does of course not entail that you shouldn't display the artworks in the first place, even if your displaying them 'enables' your visitors to take photographs of the artworks.

At first sight, then, it seems as if the pessimist does not need to take up a position that delegitimises every instance of giving out moral testimony—and this does indeed seem desirable for the pessimist. After all, such a view would be strongly revisionist. However, I worry that the pessimist cannot uphold this position all the way. After all, whatever renders reliance on moral testimony illegitimate—a lack of moral understanding, affective attitudes, autonomy or authenticity—will plausibly also be threatened by giving out moral testimony. Giving out moral testimony, after all, will also result in a lack of the important goods outlined by the pessimist. If we care about moral understanding, affective attitudes, autonomy or authenticity, moral testimony should be given out only when it does not threaten these goods.

This, however, leaves the pessimist about moral testimony in an awkward position: if giving out moral testimony is only legitimate when there is little or no risk that it will be relied upon, it is testimony only in a shallow sense. Indeed, since testimony is often understood to involve uptake as one constitutive part, whatever the pessimist declares legitimate might not even qualify as testimony in a proper sense. Despite its initial appearance, then, acknowledging not only the causal, but also the morally relevant role of the testifier in testimony-

giving commits the pessimist to an implausibly strong view—namely, that all instances of giving out moral testimony (where this is some chance of uptake) are illegitimate.

The alternative line for the pessimist is to go all in on the analogy of the art gallery and hold that pessimism about moral testimony really just means that *reliance* on moral testimony is illegitimate, with nothing being said about testifiers at all, thus absolving them of all responsibility. It is difficult to see, however, how a more nuanced position could be upheld by the pessimist.⁴

Let's pause for a moment to take stock. I suggested to understand testifying itself as an action, an action that is subject to blame and praise. Who has the relevant standing to blame a testifier when things go wrong? All of those affected—usually, at least the agent testified to, but often (at least in moral matters) also those on the receiving end of the misguided advice. I further argued that any plausible position about moral testimony should allow for a nuanced evaluation of testifiers (and advisees), that pessimism has difficulty accommodating.

6.3.3 Moral Vulnerability

Here is an interesting upshot of the proposed position: if there are more potential loci of blame in cases of moral testimony, then there can be more blame when things go wrong compared with cases without such reliance on moral testimony. This might seem bothersome (it did, at first, to me). I take it that the reason this might seem worrisome is a tacit assumption that every wrong comes with a limited amount of blame—and if a wrong is committed by two agents, the blame should be more or less split between the two.⁵ But I don't think this can be right and indeed, the principle does not seem that appealing on closer inspection. After all, seeking out testimony is usually done in an attempt to get things right. And in doing so, you are involving another person in your moral inquiry thus allowing for more points where the wrong could have been prevented (by you, by your testifier). Since neither you nor your testifier did prevent the wrong, it should not be too surprising that there can be more total blame in the present scenario.

This consideration as well as the optimist's stance that the testifier, too, can be blamed if things go wrong might seem to render testifying (at least in moral matters) unnecessarily risky. Not only can you be blamed for giving out

⁴One proposal might be that only testifiers who *manipulate* their advisee into relying on the moral testimony are blameworthy (i.e. cases where the advisee does not satisfy the requirements for moral responsibility). Here, it seems that we can legitimately blame them for causing someone to break the norm pessimism prescribes. This might be one of the ways for the pessimist to defend their position here: at least some cases where we would want to blame a testifier might be of this nature—think in particular of political figures and their respective audiences. Still, this would require an additional argument, going beyond the scope of this chapter.

⁵Perhaps this idea can be salvaged if we conceive of testimony giving as an entirely independent action, thus allowing for more blame in the present case, essentially because more (wrong) actions are taking place.

wrong testimony, you can also be blamed for things outside of your control, such as your advisees acting on your advice. By testifying, if you will, you become morally vulnerable.⁶

I do not find this result implausible but this may be due to my other convictions surrounding moral testimony. When sharing your moral beliefs, you are not (at least under normal circumstances), sharing them in a vacuum. You are sharing them where other people will respond to them—by criticising them, by changing their own views, or, as we have discussed here, by acting on them. At first, I was surprised to find that e.g. my conciliatory commitments in cases of moral disagreement I outlined in chapter 1 cohere so nicely with my optimism about moral testimony and the view of blame in cases of moral testimony I defend here (in particular, the view that the testifier receives at least some blame). But it shouldn't really be that surprising: once we take the idea that people can legitimately rely on other people in the moral domain seriously, it seems almost natural to think that the resulting blame (and praise) will be distributed. A similar coherence is to be expected for steadfast positions: if you think that we should not change our moral views in response to disagreements with others, it is not surprising to think that, in general, reliance on others is not the best strategy going forward and, because we should not change our views/actions in this way, why should anyone merely sharing their views be held responsible for 'our' wrongdoings?

But aside from theoretical commitments, taking testifying in moral matters to render you in some sense vulnerable delivers the correct verdict in a whole range of real-world cases. How is it that politicians or social media celebrities have to be particularly careful when sharing their convictions—well, because these convictions influence many people such that they might change their beliefs in response to it or even act on those convictions. And in sharing your beliefs, you are at least partially taking responsibility for that. The greater your audience, the more care you should take to avoid giving out bad (moral) testimony.⁷ On the other hand, how is it that sharing your moral beliefs in an intimate setting is, on its own, often less risky in the sense outlined? Part of the reason is that they do not influence as many people. In short, the setting of a particular instance of moral testimony matters for its evaluation in terms of blame. But this is not a special characteristic of moral testimony—it is rather a general phenomenon in our moral life.

6.4 What Kind of Inadequate Care?

So far, I argued for a picture of blame and praise when it comes to moral testimony that is perhaps surprisingly close to our usual practices of blame and

⁶At the same time, however, this position also entails that there are more loci of praise if things go right—perhaps easing some further worries regarding the riskiness of giving out moral testimony.

⁷Even Republican Senator Mitch McConnell could agree that Donald Trump was “practically and morally responsible” (McConnell (2021)) for the attack on the United States Capitol January, 6th, 2021.

praise: what matters is whether adequate care was taken, and that goes both for the agent seeking out moral testimony *and* the agent testifying.⁸

But what kind of blame are we talking about? Bracketing moral testimony for a moment, the answer seems straightforward. The agent acting wrongly because they are not sensitive to the moral features of the situation shows inadequate *moral* care. End of story. For at least some cases in which reliance on moral testimony comes into play, however, this cannot be the full story. Sure enough, the agent seeking out moral testimony might show inadequate care in this way, therefore seeking out advice. But in choosing whom to go to for advice, if they choose wrongly, do they still show inadequate *moral* care? And what about the testifying agent? It might seem tempting to think that they are to be blamed in a rather epistemic way—for their mistaken belief—as briefly touched upon in the previous section.

Again, I suggest a rather level-headed analysis of our case: sure, there is an epistemic dimension to the blame both of our agents receive. However, this is not unique to cases of reliance on moral testimony. Really, as soon as an agent starts actively deliberating about a moral choice, considering how well they are positioned to respond to it, weighing reasons in favour and against, etc., they are engaging in activities that would leave them open to a rather epistemically-flavoured kind of blame even if they did not end up relying on moral testimony. I suggest that there is little use in differentiating different kinds of blame here—because the inadequate care shown by agents will always be towards the moral features of the situation, but their deliberation about it ensures that the blame is not purely moral, but rather a mix of moral and epistemic factors. It is *both* moral and epistemic blame the agents receive. Indeed, even the agent who, as seems sometimes to be presupposed in the literature, reacts directly to the moral features of a situation can be subject to an epistemic kind of blame: perhaps they *should* have thought about the matter instead of simply responding to what they took to be the morally salient features.

Blame allocated to agents relying on or giving out moral testimony, then, is not different in flavour to the blame allocated to agents acting ‘on their own’. It is just that certain dimensions of blame become more obvious once we take moral testimony, and thus active deliberation, into account.

6.5 Conclusion

Following my discussion here, we end up with a picture on which moral testimony, against its initial appearance, does not receive a special treatment when it comes to blame. Both the agent seeking out advice and the testifier can be

⁸In his discussion of solidarity and moral responsibility, Kolers similarly argues that an account that puts full responsibility on only one side is ultimately inadequate, instead suggesting two models for sharing responsibility between the testifying agent and the agent acting on the basis of the testimony (cf. Kolers (2016), 162–165). Like me, Kolers does not identify one of these models as the perfect option, suggesting instead that contextual factors will influence which model to apply.

subject to blame, and the potential victim has the standing to blame both the advisor and the testifier. From the perspective of the testifier, this means that sharing your moral beliefs, especially when you can be reasonably sure that they might be acted upon, is risky: you may take some responsibility for the things that go wrong—without thereby removing that same responsibility from the agents acting on your advice.

Considering blame allocation in cases of moral testimony also helps us shed light on some questions which's answer we might have taken for granted before. It helps us to see the different dimensions in which blame is allocated—be it in cases with or without reliance on moral testimony. More importantly, however, we can now see the effects of relying on moral testimony when it comes to blame: relying on moral testimony does neither revoke your responsibility for your actions, nor is it particularly bad to rely on moral testimony. Likewise, giving out moral testimony is not subject to special sorts of blame—but it is not risk-free either. Moral testimony, it turns out, is much like other resources in moral inquiry, good when used well and bad when used badly.

I have argued that all of these complications can only be appreciated from the perspective of optimism about moral testimony. If there is something to these considerations, this seems to be one reason to prefer optimism over pessimism about moral testimony. Finally, from the perspective of the optimist, my discussion addresses one key worry about reliance on moral testimony: relying on moral testimony does not shift responsibility away from the agent deferring—to the contrary: the deferring agent is taking responsibility both for acting *and* for deferring. Optimism about moral testimony is thus by no means 'the easy way out' for an uncertain agent. Deferring rightly is difficult, and being optimistic about moral testimony does not change this.

Concluding Remarks

In the introduction to this dissertation, I noted that the chapters herein are all self-standing and, to some extent, independent of each other. Altogether, however, they not only cohere well in that they push us towards a more cooperative model of moral inquiry, but they do so by *interacting* with each other.

I started out by arguing for a conciliatory position when it comes to moral matters. That is, we should not form our moral beliefs in isolation—we should consider more than just our evaluation of moral matters. This opened the door to the question whether this might not be illegitimately surrendering agency. I argued that listening to others and indeed sometimes taking their word for it, is not illegitimate in that way. However, my rejection of pessimism about moral testimony is less principled than is common in the literature: I tried to carve out a theoretical space that acknowledges the concerns of the pessimist, but does not take them to defeat the legitimacy of reliance on moral testimony. I then discussed how we should incorporate such testimony into our own thoughts, arguing that the moral domain is in some sense special because all of us, whether we like it or not, are moral agents. Thus, there is a default aim of moral understanding. Emphasising the earlier point, however, this does not entail that we should always consider matters only for ourselves—many times, there will be reasons to give up on the default aim and opt for mere knowledge instead. This in turn raised the question whom we should listen to and though I did not provide an answer (pointing out the individuals/groups that are to be listened to, as it were), I don't think this is much of a problem—because expertise in moral matters will always be issue-specific. Whom to listen to simply cannot be determined in the theoretical, but must always come out of practice. Finally, I discussed the question whether we can be blameworthy for our beliefs/actions when we rely on moral testimony (yes, we can) and outlined that this is in part because that just is what it means to take moral advice seriously as a legitimate means of settling moral questions.

In some ways, my proposal here might seem to break with the traditional conception of moral inquiry—instead of reflecting mostly on our own, I suggest that others' moral beliefs are a similarly important resource in settling moral questions. In other ways, however, I don't think that the proposal is as revolutionary. As I emphasise throughout the dissertation, I take it that most people (though perhaps not most philosophers) take others' moral testimony as a useful resource in settling moral questions already anyways. Keeping this in

mind, my discussion might just as well be read as a justification of very common sense beliefs about how to conduct inquiry into moral matters. At any rate, the focus on moral testimony as a source of moral knowledge should not mislead us to think that it is the only or indeed a special way of coming to know in moral matters—my focus on moral testimony is rather due to the fact that it has often been neglected in the past literature. I think moral testimony is but one resource in moral inquiry of many, I don't think that we should defer all the time—it's just that deferring is oftentimes permissible (and sometimes required). Overall, my proposal is to understand moral inquiry as in many respects similar to non-moral inquiry: on the picture I suggest, it is the *content* that makes moral inquiry special, not our *methods* of inquiring into that content.

If all goes well in a philosophical inquiry, the end of a research project does not come with an answer to all questions. Quite the reverse: more often than not, you are left with more questions at the end of a research project than you started out with. This dissertation is no exception to this rule. I want to point out two of those newly opened questions.

First, the question whom to defer to—or, as I hope the arguments in chapter 5 made clear, the questions of whom to defer to. How specific must we think of a topic in order to answer this question? And, once we have determined the answer to this question, what are ways to identify those with a privileged standpoint on the matter? There is a long tradition we can draw on to answer these questions and I firmly believe that a full answer to them requires significant engagement not just with 'morality' as such, but with the history and context of the issue at question. In short, the answer will not be settled in a paper, it will be settled in reality.

A second question that features less prominently in the dissertation, but challenges a presupposition throughout it, concerns moral mistakes. What, if any, is the value of moral mistakes? Throughout much of the dissertation, I held fixed the idea that we want to get things right—and moral testimony is one of the resources getting us there. I wonder, however, whether there might not be a reason to think that getting things wrong might sometimes be valuable in its own right. We certainly think so for children and/or matters of little significant, but much of this can be explained via instrumental considerations: ultimately, our hope is that the agent in question gets things right—we just think that getting it wrong on this occasion might serve that goal. But perhaps this isn't the only value in moral mistakes, perhaps they can be valuable in their own right. I don't bring this up to challenge the value of just moral testimony (after all, if moral mistakes can be valuable in their own right, then so can moral mistakes based on moral testimony), but rather as a more general question as to the values in moral inquiry.

This brings me to reiterate one final point: reading the material here might make it seem as if moral testimony is special—after all, I devoted an entire dissertation to it. But I don't think so. In fact, I think moral testimony is just one resource amongst many. The novel considerations the discussion around moral testimony brings up are in fact considerations that should, for the most

part, matter also when discussing non-testimony based moral beliefs and inquiry. Really, moral testimony is special only insofar as it happens to be a resource that has been devalued more than others in moral philosophy, and I hope to have made a convincing case that it has been devalued unjustifiedly. It is okay, and often good, to *listen* to others. And it is okay, and often good, to *trust* others—also when it comes to moral matters.

Bibliography

- Elizabeth Anderson. Feminist Epistemology and Philosophy of Science. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, Spring 2020 edition, 2020.
- David Archard. Why Moral Philosophers Are Not and Should Not Be Moral Experts. *Bioethics*, 25(3):119–127, 2011.
- Nomy Arpaly. Moral Worth. *Journal of Philosophy*, 99(5):223, 2002.
- Rima Basu. The Wrongs of Racist Beliefs. *Philosophical Studies*, 176(9):2497–2515, 2019.
- Nora Berenstain. Epistemic Exploitation. *Ergo: An Open Access Journal of Philosophy*, 3:569–590, 2016.
- Monica Biernat and Diane Kobrynowicz. Gender- and race-based standards of competence: lower minimum standards but higher ability standards for devalued groups. *Journal of Personality and Social Psychology*, 72:544–557, 1997.
- Johan Brännmark. Moral Disunitarianism. *Philosophical Quarterly*, 66(264):481–499, 2016.
- Michael Bratman. *Intention, Plans, and Practical Reason*. Cambridge: Cambridge, MA: Harvard University Press, 1987.
- George Bridges and Sara Steen. Racial disparities in official assessments of juvenile offenders: Attributional stereotypes as mediating mechanisms. *American Sociological Review*, 63:554–570, 1998.
- Jessica Brown. Assertion and Practical Reasoning: Common or Divergent Epistemic Standards? *Philosophy and Phenomenological Research*, 84(1):123–157, 2012.
- Tyler Burge. Content Preservation. *Philosophical Review*, 102(4):457–488, 1993.
- Laura Frances Callahan. Moral Testimony: A Re-Conceived Understanding Explanation. *Philosophical Quarterly*, 68(272):437–459, 2018.

- Yuri Cath. Knowing What It is Like and Testimony. *Australasian Journal of Philosophy*, 97(1):105–120, 2019.
- Michael Cholbi. Moral Expertise and the Credentials Problem. *Ethical Theory and Moral Practice*, 10(4):323–334, 2007.
- Michael Cholbi. Why Moral Expertise Needs Moral Theory. In Jamie Carlin Watson and Laura K. Guidry-Grimes, editors, *Moral Expertise: New Essays from Theoretical and Clinical Bioethics*, pages 71–86. Springer International Publishing, 2018.
- David Christensen. Epistemology of Disagreement: The Good News. *Philosophical Review*, 116(2):187–217, 2007.
- David Christensen. Conciliation, Uniqueness and Rational Toxicity. *Noûs*, 50(3):584–603, 2016.
- Stewart Cohen. Basic Knowledge and the Problem of Easy Knowledge. *Philosophy and Phenomenological Research*, 65(2):309–329, 2002.
- Roger Crisp. Moral Testimony Pessimism: A Defence. *Aristotelian Society Supplementary Volume*, 88(1):129–143, 2014.
- Casey Doyle. Remembering what is right. *Philosophical Explorations*, 23(1):49–64, 2020.
- Julia Driver. Moral Expertise: Judgment, Practice, and Analysis. *Social Philosophy and Policy*, 30(1-2):280–296, 2013.
- Nicole Dular. Moral Testimony Under Oppression. *Journal of Social Philosophy*, 48(2):212–236, 2017.
- Adam Elga. Reflection and Disagreement. *Noûs*, 41(3):478–502, 2007.
- David Enoch. Not Just a Truthometer: Taking Oneself Seriously (but Not Too Seriously) in Cases of Peer Disagreement. *Mind*, 119(476):953–997, 2010.
- David Enoch. A Defense of Moral Deference. *Journal of Philosophy*, 111(5):229–258, 2014.
- Richard Feldman. Epistemological Puzzles About Disagreement. In Stephen Hetherington, editor, *Epistemology Futures*, pages 216–236. Oxford University Press, Oxford, 2006a.
- Richard Feldman. Reasonable Religious Disagreements. In Louise Antony, editor, *Philosophers Without Gods: Meditations on Atheism and the Secular Life*, pages 194–214. Oxford University Press, Oxford, 2006b.
- Guy Fletcher. Moral Testimony: Once More with Feeling. *Oxford Studies in Metaethics*, 11:45–73, 2016.

- Paulette C. Flore and Jelte M. Wicherts. Does stereotype threat influence performance of girls in stereotyped domains? A meta-analysis. *Journal of School Psychology*, 53(1):25–44, 2015.
- Michael Allen Fox. Vegetarianism and Planetary Health. *Ethics and the Environment*, 5(2):163–174, 2000.
- Bryan Frances and Jonathan Matheson. Disagreement. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, Winter 2019 edition, 2019.
- Paulo Freire. *Pedagogy of the Oppressed. 30th Anniversary Edition*. Continuum, New York, 2000.
- Elizabeth Fricker. Second-Hand Knowledge. *Philosophy and Phenomenological Research*, 73(3):592–618, 2006.
- Jane Friedman. The Aim of Inquiry? *Philosophy and Phenomenological Research*, Forthcoming.
- James Fritz. Conciliationism and Moral Spinelessness. *Episteme*, 15(1):101–118, 2018.
- Allan Gibbard. *Thinking How to Live*. Harvard University Press, 2003.
- Sanford Goldberg. Reductionism and the Distinctiveness of Testimonial Knowledge. In *The Epistemology of Testimony*. Oxford University Press, 2006.
- Sanford Goldberg. Putting the Norm of Assertion to Work: The Case of Testimony. In *Assertion: New Philosophical Essays*. Oxford University Press, 2011.
- Alvin I. Goldman. *Knowledge in a Social World*. Oxford University Press, Oxford, 1999.
- Alvin I. Goldman. Experts: Which Ones Should You Trust? *Philosophy and Phenomenological Research*, 63(1):85–110, 2001.
- John-Stewart Gordon. Moral Philosophers Are Moral Experts! A Reply to David Archard. *Bioethics*, 28(4):203–206, 2014.
- Daniel Groll and Jason Decker. Moral Testimony: One of These Things Is Just Like the Others. *Analytic Philosophy*, 55(1):54–74, 2014.
- Elizabeth Harman. Ethics is Hard! What Follows?, 2014. URL <http://www.princeton.edu/~eharman/Ethics%20Is%20Hard%20020714%20For%20Web.pdf>.
- Elizabeth Harman. Moral Testimony Goes Only so Far. *Oxford Studies in Agency and Responsibility*, 6:165–185, 2019.

- Amelia Hicks. Moral Hedging and Responding to Reasons. *Pacific Philosophical Quarterly*, 100(3):765–789, 2019.
- Alison Hills. Moral Testimony and Moral Epistemology. *Ethics*, 120(1):94–127, 2009.
- Alison Hills. Moral Testimony. *Philosophy Compass*, 8(6):552–559, 2013.
- Ulf Hlobil. Why You Cannot Make People Better by Telling Them What is Good. *European Journal of Philosophy*, 28(4):986–996, 2020.
- Robert Hopkins. What is Wrong with Moral Testimony? *Philosophy and Phenomenological Research*, 74(3):611–634, 2007.
- Robert J. Howell. Google Morals, Virtue, and the Asymmetry of Deference. *Noûs*, 48(3):389–415, 2014.
- Michael Huemer. *Ethical Intuitionism*. Palgrave Macmillan, Houndmills, 2005.
- Zoë Johnson King. Accidentally Doing the Right Thing. *Philosophy and Phenomenological Research*, 100(1):186–206, 2020a.
- Zoë Johnson King. Praiseworthy motivations. *Noûs*, 54(2):408–430, 2020b.
- Zoë Johnson King. Deliberation and Moral Motivation. In Russ Shafer-Landau, editor, *Oxford Studies in Metaethics, Volume 17*. Oxford University Press, 2022.
- Karen Jones. Second-Hand Moral Knowledge. *Journal of Philosophy*, 96(2):55–78, 1999.
- Thomas Kelly. Peer Disagreement and Higher Order Evidence. In Richard Feldman and Ted A. Warfield, editors, *Disagreement*. Oxford University Press, 2010.
- Christoph Kelp. Theory of Inquiry. *Philosophy and Phenomenological Research*, 103(2):359–384, 2021.
- Avery Kolers. *A Moral Theory of Solidarity*. Oxford University Press, 2016.
- Hilary Kornblith. Belief in the Face of Controversy. In Richard Feldman and Ted A. Warfield, editors, *Disagreement*, pages 29–52. Oxford University Press, Oxford, 2010.
- Christine Korsgaard. *The Sources of Normativity*. Cambridge University Press, 1996.
- Christine Korsgaard. Interacting With Animals: A Kantian Account. In Tom L. Beauchamp and R. G. Frey, editors, *Oxford Handbook of Animal Ethics*, pages 91–118. Oxford University Press, Oxford, 2012.

- C. Kunkler. *The 89th Annual Kate Kennedy Procession: College Echoes*. Kate Kennedy Club University of St Andrews, St Andrews, 2015.
- Nick Küspert. Conciliating to avoid moral scepticism. *International Journal of Philosophical Studies*, 31(3):279–300, 2023.
- Jennifer Lackey. A Justificationist View of Disagreement’s Epistemic Significance. In Adrian Haddock, Alan Millar, and Duncan Pritchard, editors, *Social Epistemology*, pages 298–325. Oxford University Press, Oxford, 2010.
- Jennifer Lackey. Experts and Peer Disagreement. In Matthew A. Benton, John Hawthorne, and Dani Rabinowitz, editors, *Knowledge, Belief, and God: New Insights in Religious Epistemology*, pages 228–245. Oxford: Oxford University Press, 2018.
- Eimar Leahy, Sean Lyons, and Richard Told. An Estimate of the Number of Vegetarians in the World. *Economic and Social Research Institute*, 340:1–31, 2010.
- Hallvard Lillehammer. Smith on Moral Fetishism. *Analysis*, 57(3):187–195, 1997.
- Hallvard Lillehammer. Moral Testimony, Moral Virtue, and the Value of Autonomy. *Aristotelian Society Supplementary Volume*, 88(1):111–127, 2014.
- Errol Lord. How to Learn About Aesthetics and Morality Through Acquaintance and Deference. *Oxford Studies in Metaethics*, 13, 2018.
- J. L. Mackie. *Ethics: Inventing right and wrong*. Penguin, Harmondsworth, 1977.
- Barry Maguire and Justin Snedegar. Normative Metaphysics for Accountants. *Philosophical Studies*, 2020.
- Julia Markovits. Acting for the Right Reasons. *Philosophical Review*, 119(2): 201–242, 2010.
- Mitch McConnell. McConnell on Impeachment: “Disgraceful Dereliction” Cannot Lead Senate to “Defy Our Own Constitutional Guardrails”, 2021. URL <https://www.mcconnell.senate.gov/public/index.cfm/pressreleases?ID=2ED4EA6B-4BEE-462C-BF1E-7972207B5F53>.
- Sarah McGrath. The Puzzle of Pure Moral Deference. *Philosophical Perspectives*, 23(1):321–344, 2009.
- Paddy Jane McShane. Moral Testimony and Moral Understanding. *Journal of Moral Philosophy*, 15(3):245–271, 2018.
- Benjamin De Mesel. Are Moral Judgements Semantically Uniform? A Wittgensteinian Approach to the Cognitivism - Non-Cognitivism Debate. In Benjamin De Mesel and Oskari Kuusela, editors, *Ethics in the Wake of Wittgenstein*, pages 126–148. New York: Routledge, 2019.

- Andreas L. Mogensen. Moral Testimony Pessimism and the Uncertain Value of Authenticity. *Philosophy and Phenomenological Research*, 95(2):261–284, 2017.
- Shyam Nair. How Do Reasons Accrue? In E. Lord and B., editors, *Weighing Reasons*, pages 58–73. Oxford University Press, New York, 2016.
- Dana Nelkin. Moral Luck. In Edward N. Zalta and Uri Nodelman, editors, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, Spring 2023 edition, 2023.
- Philip Nickel. Moral Testimony and its Authority. *Ethical Theory and Moral Practice*, 4(3):253–266, 2001.
- Joseph Raz. The Morality of Freedom. *Ethics*, 98(4):850–852, 1988.
- G. Ritchie. A town’s history is right at hand. *The Courier & Advertiser*, page 30, 2015.
- Gideon Rosen. Skepticism About Moral Responsibility. *Philosophical Perspectives*, 18(1):295–313, 2004.
- Lewis D. Ross. Is Understanding Reducible? *Inquiry: An Interdisciplinary Journal of Philosophy*, 63(2):117–135, 2020.
- Abraham Roth. Entitlement to reasons for action. In David Shoemaker, editor, *Oxford Studies in Agency and Responsibility vol. 4*, pages 75–92. Oxford University Press, 2017.
- Thomas Scanlon. *What We Owe to Each Other*. Belknap Press of Harvard University Press, 1998.
- Toni Schmader, Michael Johns, and Marchelle Barquissau. The Costs of Accepting Gender Differences: The Role of Stereotype Endorsement in Women’s Experience in the Math Domain. *Sex Roles*, 50:835–850, 2004.
- Miriam Schoenfield. Permission to Believe: Why Permissivism Is True and What It Tells Us About Irrelevant Influences on Belief. *Noûs*, 48(2):193–218, 2014.
- Andrew Sepielli. Moral Uncertainty and Fetishistic Motivation. *Philosophical Studies*, 173(11):2951–2968, 2016.
- Kieran Setiya. Does Moral Theory Corrupt Youth? *Philosophical Topics*, 38(1):205–222, 2010.
- Kieran Setiya. *Knowing Right From Wrong*. Oxford University Press, Oxford, 2012.
- Russ Shafer-Landau. *Whatever Happened to Good and Evil?* Oxford University Press, Oxford, 2003.

- Henry Sidgwick. *The Methods of Ethics. 7th Edition*. Macmillan and Company, London, 1908.
- Peter Singer. Moral Experts. *Analysis*, 32(4):115–117, 1972.
- Walter Sinnott-Armstrong and Thalia Wheatley. The Disunity of Morality and Why it Matters to Philosophy. *The Monist*, 95(3):355–377, 2012.
- Paulina Sliwa. In Defense of Moral Testimony. *Philosophical Studies*, 158(2): 175–195, 2012.
- Angela M. Smith. On Being Responsible and Holding Responsible. *The Journal of Ethics*, 11(4):465–484, 2007.
- Michael Smith. *The Moral Problem*. Blackwell, Oxford, 1994.
- Justin Snedegar. Reasons for and Reasons Against. *Philosophical Studies*, 175 (3):725–743, 2018.
- Dan Sperber and Hugo Mercier. *The Enigma of Reason*. Harvard University Press, 2017.
- Julia Staffel. Credences and suspended judgments as transitional attitudes. *Philosophical Issues*, 29(1):281–294, 2019. doi: 10.1111/phs.12154.
- Julia Staffel. Transitional attitudes and the unmooring view of higher-order evidence. *Noûs*, 57(1):238–260, 2021.
- Daniel Star. *Knowing Better: Virtue, Deliberation, and Normative Ethics*. Oxford University Press, 2015.
- Claude Steele and Joshua Aronson. Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology*, 69(5):797–811, 1995.
- Emily Tilton. “That’s Above My Paygrade”: Woke Excuses for Ignorance. *Philosophers’ Imprint*, Forthcoming.
- Katia Vavova. Moral Disagreement and Moral Skepticism. *Philosophical Perspectives*, 28(1):302–333, 2014.
- Brian Weatherson. Running Risks Morally. *Philosophical Studies*, 167(1):141–163, 2014.
- Ralph Wedgwood. The Moral Evil Demons. In Richard Feldman and Ted A. Warfield, editors, *Disagreement*. Oxford University Press, 2010.
- Roger White. Epistemic Permissiveness. *Philosophical Perspectives*, 19(1):445–459, 2005.
- Eric Wiland. Moral Testimony: Going on the Offensive. *Oxford Studies in Metaethics*, 12, 2017.

Eric Wiland. *Guided by Voices: Moral Testimony, Advice, and Forging a 'We'*. Oxford, UK: Oxford University Press, 2021.

Michael Williams. *Problems of Knowledge: A Critical Introduction to Epistemology*. Oxford University Press, Oxford, 2001.

Robert Paul Wolff. *In Defense of Anarchism*. University of California Press, 1970.

Linda Trinkaus Zagzebski. *Epistemic Authority: A Theory of Trust, Authority, and Autonomy in Belief*. Oxford University Press, 2012.