

KNOWING AND DOING, WELL AND GOOD

Abby Everett Jaques

A.B. Art & Archæology
Princeton University

*Submitted to the department of Linguistics & Philosophy
in partial fulfillment of the requirements for the degree of*

Doctor of Philosophy
at the
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2018

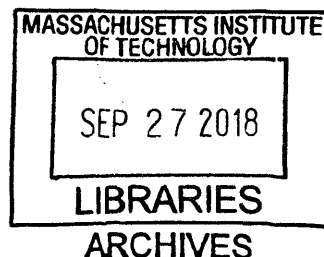
© 2018 Abby Everett Jaques. All rights reserved.

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

Signature of Author: _____ **Signature redacted**
Department of Linguistics & Philosophy
7 September 2018

Certified by: _____ **Signature redacted**
Kieran Setiya
Professor of Philosophy
Thesis Supervisor

Accepted by: _____ **Signature redacted**
Bradford Skow
Laurance S. Rockefeller Professor of Philosophy
Chair of the Committee on Graduate Students



KNOWING AND DOING, WELL AND GOOD

Abby Everett Jaques

*Submitted to the department of Linguistics & Philosophy
on 7 September 2018 in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Philosophy.*

ABSTRACT

My dissertation explores the relationship between intentional action and knowledge. I argue that understanding this relationship is not only of central importance to action theory; it is also a means to progress on questions about the nature of knowledge, the mechanisms of oppression, and the foundations of ethics.

The opening chapter argues for my view of the nature of intentional action. I show that it is distinguished from nearby phenomena by an aim of control—and that this control turns out to be Anscombean *practical knowledge*, the special knowledge an agent can have of what she is doing, how, and why.

The second chapter considers how my view about the knowledge–action relationship differs from those advocated by ‘shifty epistemologists’—theorists who claim that what you know depends on practical factors like what’s at stake for you. I argue that my view undermines the motivation for this claim and may debunk it.

The third chapter presents a new way of understanding epistemic injustice and describes how epistemic injustice (thus understood) interacts with action’s constitutive aim of practical knowledge to cause *shackling*—a distinctive dilemma faced by marginalized agents that both manifests and constitutes oppression.

The fourth chapter shows how my view of the nature of intentional action entails a new kind of constitutivism about practical reason. I raise a worry that this form of constitutivism threatens the existence and/or generality of moral reasons before suggesting some possible ways out.

Thesis supervisor: Kieran Setiya
Title: Professor of Philosophy

ACKNOWLEDGEMENTS

I am deeply indebted and wildly grateful for the generous and helpful comments and support I received from my committee, colleagues, and friends.

My advisor, Kieran Setiya, spent untold hours discussing every part of this thesis; he also taught me to read Anscombe, which is to say that without him, this work would not exist. My other committee members—Sally Haslanger, Jason Stanley, and Roger White—contributed not only essential feedback but much more. Sally made me want to be a philosopher, and her thought runs through every part of my own. Jason’s work on the relationship between knowledge and action and the political importance of that relationship provided an animating influence on everything here. And Roger was the one who first introduced me to thinking about knowledge and action together, on a fateful day in my first year at MIT. I am constantly inspired by the mentorship, teaching, and writing of these four philosophers.

Profound thanks to Nancy Bauer, Avner Baz, David Boylan, Alex Byrne, Kevin Dorst, Nykki Dular, Cosmo Grant, Lyndal Grant, Caspar Hare, Sam Hesni, Matthias Jenny, Erin Kelly, Justin Khoo, Ari Koslow, Rae Langton, Rose Lenehan, Colin Marshall, Rachel McKinney, Lionel McPherson, Christiana Olfert, Sarah Paul, Milo Phillips-Brown, Kevin Richardson, Wendy Salkin, Jeff Sebo, Tamar Schapiro, Ginger Schultheis, George Smith, Judy Thomson, Yafeng Wang, and Tim Williamson for providing indispensable help and encouragement.

Parts of this work, in various forms and at various stage of development, benefited greatly from being presented at conferences. Thanks to the or-

ganizers and audiences at the 2017 Wisconsin Metaethics Workshop (Mad-Meta); the 2017 Edinburgh Law & Metaphysics Workshop; the 2016 New Methods of Ethics conference at the University of Birmingham; the 2016 Rocky Mountain Ethics Congress (RoME), and to Iskra Fileva, for very useful written comments on the work presented there; a 2016 session of the Workshop on Gender and Philosophy; and the 2015 MITing of the Minds conference.

Finally, I want to thank my spouse, John. It is not easy to live with a dissertating human. There are tens of thousands of words in this thesis, and I only had the chance to produce them because of his unfailing love and support.

CONTENTS

1	WHERE THERE'S A WILL, THERE'S A WAY OF KNOWING	9
1.1	Anscombe's thought	11
1.2	Counterexamples	14
1.3	Saving the Anscombean thought: previous proposals	15
1.4	My proposal: Mind control	23
1.5	Control as practical knowledge	24
1.6	Revisiting Davidson's carbon copier case	35
1.7	Conclusion	41
2	KNOWING WHAT YOU'RE DOING WITH SHIFTY EPISTEMOLOGY	45
2.1	Two kinds of cases	48
2.1.1	The criticism cases	50
2.1.2	The encroachment cases	53
2.2	Intentional action's aim: practical knowledge	55
2.2.1	Saving the Anscombean thought	58
2.2.2	A clarification and a caveat	61
2.3	A different flavor of knowledge norm	62
2.4	Comparing results in cases	70
2.4.1	HSP versus KNOWN REASONS in criticism cases	70
2.4.2	HSP versus KNOWN REASONS in encroachment cases	72
2.4.3	A caveat: Not just cases	79
2.5	Conclusion	80

3	SHACKLING	85
3.1	What is epistemic injustice?	86
3.1.1	The epistemic economy	87
3.1.2	Testimonial injustice, revisited	94
3.1.3	Hermeneutical injustice, revisited	96
3.1.4	The key differences in my picture compared to Fricker's	99
3.2	What is intentional action?	101
3.3	Shackling	106
3.4	Cases	112
3.4.1	Shackling via hermeneutical injustice	113
3.4.2	Shackling via hermeneutical injustice, part 2	126
3.4.3	Shackling via credal injustice	128
3.4.4	A complication: combined cases	131
3.4.5	A final complication: defective concepts	132
3.5	Conclusion	134
4	KNOWING AND DOING, WELL AND GOOD	141
4.1	I hope you know what you're doing	142
4.2	Practical knowledge constitutivism	147
4.3	Who's afraid of practical knowledge constitutivism?	153
4.4	One way out: Oomphy concepts	157
4.5	Another way out: Knowing all the way	163
4.6	Better together	170
4.7	Lots of lingering questions	171
4.7.1	Down with imperialism	172
4.7.2	Extra! Extra!	173
4.7.3	Naive saints	174
4.8	Conclusion	176

1 WHERE THERE'S A WILL, THERE'S A WAY OF KNOWING

Consider two agents—God, and Wile E. Coyote. God is a *perfect* agent: omnipotent, omniscient, omnibenevolent. Wile E. Coyote is a mess: failing, fallible, and at fault. This paper isn't about religion or cartoons, but it's useful to think about both perfect and entirely imperfect agents as we get started.

I'm going to argue for a view on which God's omnipotence is bound up with her omniscience, and vice versa—and on which Wile E. Coyote's failures of knowledge and action are likewise intertwined. These things follow from a line of thought unspooled in reflecting on a famous—and contested—claim of Anscombe's. She said that when you act intentionally, you know what you're doing, in a special way. This isn't right (as Davidson showed), but it's onto something important. There really is a crucial, constitutive link between knowing and doing that explains why God's omniscience and omnipotence go together, and why poor Wile E. is never going to catch the Road Runner. The principle I propose says that intentional action constitutively aims at control—and the control in question turns out to be practical knowledge, in Anscombe's sense: the special knowledge an agent can have of what she is doing, and how, and why.

To start, I'll review the relevant bit of Anscombe and the counterexamples that have been taken to be fatal to her way of thinking, sketch some previous replies on Anscombe's behalf, and explain why I think they don't quite give us what we should want. That's §§1.1–1.3. Next, I'll argue for the principle

I call MIND CONTROL, which says that whatever else it may be, intentional action is an exercise of agency such that something the agent has in mind is supposed not only to match up with something that happens in the world, but also to *guide and secure* that happening. I'll show that these conditions are needed to distinguish intentional actions from nearby phenomena—mere behaviors, things that befall agents, instances of deviant causation, etc.—and that what they capture is the agential control that is distinctive of intentional action. That's §1.4.

The heart of the paper is §1.5. There I'll show that MIND CONTROL turns out to have several important upshots. First, control comes by degrees, so what we want is a normative, rather than a necessary, condition. Second, the securing condition means that what is required is for mind and world to align *and not as a matter of luck*—and, third, that means that what is required is *knowledge*. Finally, the guidance requirement tells us what the agent must know, and the answer is that she must know *what she is doing, and why, and how*. All together, this means that to say that it is constitutive of intentional action to aim at control is to say it is constitutive of intentional action to aim at Anscombean practical knowledge. With all that in hand, in §1.6 I walk through the most famous Davidsonian counterexample to Anscombe's thought, the doubtful carbon copier, and show how my view makes good sense of the case.

Thinking about action in the way I develop in this paper doesn't just help make Anscombe's thought less outlandish. It poses a new challenge to views—such as, most famously, Michael Bratman's—on which intentions are non-cognitive, distinctively practical attitudes. The challenge arises because my argument brings out a deep connection between intentional action and knowledge that is not obviously explicable on a view that takes intentions to be non-cognitive. There has been something of a stalemate between cognitivists about intention and their opponents; I take it this paper shifts the balance in that debate.

Finally, because the arguments of this paper weave together strands from

action theory, epistemology, and and beyond, the resulting view has upshots in several other areas of philosophy. I will briefly mention some of them in §1.7, though full discussion will require a great deal of further work.

1.1 ANSCOMBE'S THOUGHT

Anscombe (1963, §§5–7) saw the connection between knowing and doing. She said when you act intentionally, you *know what you are doing*, and in a special way—you needn't look to see, and you needn't reason your way to the conclusion. You know *without observation* or inference.¹ Call this:

Anscombe's Thought: When you act intentionally, you know what you are doing in a special way.

There is something intuitive here; we can build up. First, it seems right to say that if you are utterly convinced that you are not ϕ -ing, then you cannot be ϕ -ing intentionally.² Second, as Anscombe explains, it also seems right to say that one is not doing intentionally what one is simply *unaware* that one is doing.

She invites us to consider a test: if you ask me why I am typing so loudly, and I reply (truthfully) that I was not aware that I was doing so, then my typing loudly was not intentional.³ On the other hand, if you ask me why I am typing, I will reply that I am writing a paper: my typing is intentional, even if my typing loudly is not. To answer the 'why?' question is to accept that the action about which the question is asked is intentional under the description used in the question, and to give a reason for which the action is done; to say instead that one is unaware of ϕ -ing under some description

1. See Anscombe 1963, e.g. 13–15, 49–50. Setiya (2007, 24; 2008; 2011) emphasizes the bar on inference, also citing Hampshire (1959, 70).

2. Setiya (2008) makes this point.

3. Anscombe 1963, 9.

is to refuse application of the question, and thereby indicate that the action is not intentional under that description.⁴

This, then, is the kind of knowledge Anscombe says an agent has of what she is doing: she knows that she is ϕ -ing under the description used in the ‘why?’ question, and she has an answer to the ‘why?’ question, which is her reason. Together, these parts enable her answer. As Anscombe (1963, 9) says, the sense of ‘why?’ at work in the test “is of course that in which the answer, if positive, gives a reason for acting.”

What kind of things might one offer as a reason for acting in this way? There are a few basic types. Again, suppose you ask me why I am (still, doggedly, but hopefully not quite so loudly) typing. I might reply in any of the following ways:

- (1) I’m writing Chapter 1.
- (2) In order to finish this draft.
- (3) My deadline is tomorrow.
- (4) I want to finish before teatime.

These are paradigmatic positive answers to the ‘Why?’ question, and so mark my action as intentional. Note their variety. One says that my typing (partly) constitutes another action; one points to an aim I have; one asserts a fact about my circumstances; one cites a desire of mine. All these kinds of things, and more, can be reasons in the sense Anscombe is after.

On the other hand, there are some rather different kinds of reply I might have offered instead:

- (4) No particular reason.

4. Anscombe 1963, 11.

(5) I wasn't aware I was doing that.⁵

(6) I seem to be having a curiously elaborate spasm.

Anscombe says that (4) still counts as the kind of reply that indicates intentional action, even though it is a denial that there is a reason for what you're doing.⁶ This is a somewhat odd kind of case, and Davidson (2001, 6) and others have disagreed with Anscombe about the possibility of acting intentionally for no particular reason. But for her, *intentional action* and *action performed for reasons* are not equivalent. While actions performed for reasons will be intentional, the reverse can fail to be the case. We can get a sense of why Anscombe thinks so, and thus groups (4) with (1)–(3), by contrasting (4) with (5).⁷

If, when you're asked why you're ϕ -ing, you say that you have no reason, you are still accepting that the question is apt. But if you say that you weren't aware of the ϕ -ing, you are *refusing application* of the question. That's the case in (5), and for Anscombe a (5) kind of answer marks actions that are not intentional (or, at least, descriptions under which an action is not intentional).

What about (6)? It looks like a positive answer, but a spasm is exactly the kind of thing that we want to class as unintentional. I won't recount all of Anscombe's rather complex discussion of this kind of case here; I'll just sketch the additional feature she adds to her account in order to deal with it.

When I give an answer like (6), I am able to answer the question; I know that I am ϕ -ing (typing) because p (I am having a curiously elaborate spasm). So

5. Though obviously in the case of typing this would require some explaining!

6. See her §17.

7. In what follows, I'll leave open the possibility of intentional action for no particular reason.

structurally, this looks right, and it seems less strange in some ways than (5). But if Anscombe is committed to spasms' counting as intentional actions, her account is a non-starter—and nothing she has said so far gives her the resources to bar spasms from the class of intentional actions.

And so Anscombe (1963, §§8–9) asks that we notice a further distinction. Suppose I do know what I'm doing in (6) as well as (1)–(4). Still, I don't seem to know in the same way. In the spasm case, I as it were *find myself* typing; we might say that I infer the spasm from observing my movements and perhaps noting the lack of the kind of answer given in (1)–(4). As Anscombe puts it, I require (empirical) *evidence* for (6), but in (1)–(4) I know what I am doing *without observation* (or inference).

There is still much that is puzzling about this picture. In what follows, I will address some worries, offer some clarifications, and ultimately argue for an amendment. For now, I will merely summarize. Knowledge without observation or inference is necessary (though not sufficient) for intentional action, according to Anscombe. When an action is intentional, she says, the 'why?' question is given application; the ability to answer the question involves the agent's knowing what she is doing in a special way, viz., without observation or inference. Thus intentional action has essentially to do with knowledge. There is more to say—in particular, about the distinctively practical nature of this knowledge—but this rough gloss should suffice for our purposes at this point.

1.2 COUNTEREXAMPLES

Here's a problem, though: Anscombe's thought, as stated, is just false.

It seems that I can, for example, ϕ intentionally, believe that I am ϕ -ing, but fail to know it, even if I will succeed. Suppose I need to get to class by 2 p.m. I look at my watch, and it says 1:47, so I pick up my things and walk out the door. If you ask me what I'm doing and why, I'll say I'm heading to class because it's time to go. But if my watch stopped at 1:47 a.m., and it's just a fluke that when I looked at it it was in fact 1:47 p.m., we can think

of this as a case where I believe, but don't know, that I'm leaving on time to go to class, because I'm Gettiered—and yet it certainly seems that I will have left for class on time intentionally.

What's more, as Davidson (2001, 91–92) points out, it is perfectly possible to act intentionally without even believing that one is doing what one intends—and so without knowing that one is doing it, even if one succeeds. He describes someone trying to make ten carbon copies at once (writing heavily, pressing down hard on the stack of paper and carbons) while not at all confident of the outcome, indeed, in a state of extreme doubt that he is succeeding. It may be that the writer will in fact succeed in making ten copies; if so, he will have done so intentionally, yet without knowing *or believing* himself to be doing it. This kind of case has been widely thought to pose a devastating challenge to Anscombe's point.⁸

If one can fail even to believe that one is ϕ -ing, and yet still end up ϕ -ing intentionally, it is tempting to think there is no significant connection between intending and knowing and doing after all; Anscombe's thought seems on the brink of death by counterexample. Is there any hope?

1.3 SAVING THE ANSCOMBEAN THOUGHT: PREVIOUS PROPOSALS

A few people have thought so. The two main efforts to save Anscombe's thought have come from Setiya and Thompson; I'll start with Thompson.

On Thompson's (2008, 2011) way of thinking, the counterexamples basically aren't counterexamples at all: because we properly describe intentional actions via the progressive, the weakness of that form means there is no great obstacle to knowing what one is doing. He is right about the progressive—it is certainly true that I am writing this paper at this moment, even if I will never finish it. And we might think I will still be writing this paper when

8. Though see below.

I am asleep tonight, since I will neither have finished nor have abandoned the task. It doesn't matter that I may never succeed in having ϕ -ed: it can be true to say of me that I am writing this paper even if a moment from now my laptop and I will be crushed by an asteroid, and so it will never be the case that I have written it. That is what it means for the progressive to be weak; it carries no implication of success. Pushing that thought further, Thompson (2008) says that in fact as soon as I form the intention to ϕ , I am what he calls *IMP to ϕ* , where being *IMP to ϕ* is just to be, in perhaps the thinnest way possible, in the process of ϕ -ing. So there is a sense in which, as he sees it, for me to intend to ϕ , even if I have as yet taken no steps, is for me to be, already, ϕ -ing.⁹ Thus if you and I are in the grocery store on a trip in which you will soon accidentally take skim milk from the shelf, instead of the whole milk that is on the list, you are nonetheless buying the items on the list, because that is what you intend, even if it will not be the case that you have done so when we are through.¹⁰

Thompson (2011, 209) emphasizes the idea that practical knowledge is *self-knowledge*, like one's knowledge that one is in pain:

My thoughts and pains are matters for self-consciousness, only as long as I have them, as long as they are present. It is the same with practical knowledge, and thus since the present in this case must be imperfective, there is practical knowledge only when the thing is precisely NOT done, not PAST; there is more to come, something thing is missing, and the H-bomb may hit before it does. My so-called knowledge of my intentional action in truth exists only and precisely when there is no action, but only something I am doing.

9. The strength of this claim varies between his (2008) and his (2011).

10. It even looks like on Thompson's view you can perhaps be buying three bottles of invisibility potion, if that is what you have sincerely set out to do, until the moment you give up on doing so.

If Thompson is right, then knowing what one is doing is much easier than we might have supposed, because intending is sufficient for being in progress, and so intending is sufficient for doing. When we worry about cases of picking up the wrong carton in the dairy aisle, or failing to bring our doings to fruition in general, we are in some sense changing the subject.

So far, I am inclined to think this raises as many questions as it answers. But since Thompson (2011, 209–10) addresses the carbon copier case specifically, it may help to consider what he says about it. Thompson reasonably notes that there is something peculiar about Davidson’s version of the case: in such a scenario one would naturally *check* after a few strokes to see if the marks were making it all the way to the bottom. Thompson says that therefore we need to distinguish a case that proceeds in this more natural way, with the copy-maker checking as he goes, splitting the pile and doing more passes as needed, from a one-off, no-checking version like Davidson’s.

Thompson says that in the more natural, checking-as-he-goes scenario, the carbon copier knows he is making ten copies, all along. In the one-off, no-checking version, he doesn’t know, and so it isn’t actually right, Thompson says, to say that if he succeeds, he has made his ten copies intentionally. The one-off copier’s copy-making is like buying a lottery ticket, and in both cases lucky success does not count as intentional. As Gibbons (2001, 587–88) says:

Consider the following three lottery stories. In the first, Cindy buys a lottery ticket, knowing her chances are a million to one, and she wins. In this case, with a fair lottery, Cindy’s winning is too accidental, or too lucky to count as intentional. In the second case, Cindy mistakenly believes someone rigged the lottery in her favor. She believes, on the basis of this, that if she buys a ticket, she will win. She buys the ticket and wins. So her belief about winning is true. She even has a justified true belief. But her winning is just as accidental and just as lucky as it was in the first case. So her winning is not intentional. What is missing? It looks like knowledge. In the third case, Cindy knows the lottery is rigged and knows that if she buys a ticket she will win. In this

case, Cindy intentionally wins the lottery.

In what follows, I will have a good deal to say about the relationship between luck, intentional action, and practical knowledge. (I will not endorse the line Gibbons pushes above, but my view will make it clear why he is tempted to think that way.) But for now, I want to note a couple of worries that have been voiced about Thompson's view. First, Thompson's view does not yield especially plausible results in cases of intentions for the future, which is desirable if we want a unified account of intention (as Thompson says he does): suppose we grant that I can be, and can know that I am, writing this paper, although I will never have written it, because of an asteroid strike; it seems much less plausible to say that upon forming the intention to, say, stand perfectly still at 6:14 p.m. tomorrow, I both am doing so and know it.¹¹ Second, the stark bifurcation of Thompson's account of the carbon copier case can seem unsatisfying.

Those are standard worries. But what I want to focus on is something else: In essence, Thompson's way of vindicating Anscombe's thought amounts to trivializing it. Consider it from the agent's point of view. Thompson says that she knows what she's doing, even in the problem cases, because what she is doing is what she intends, even if she will never have done what she intends. But why might an agent *want* to know what she is doing, on this way of thinking?

When we ask why an agent would be interested in knowing that (or whether) she is ϕ -ing, the natural thought is that she is interested in knowing that she is on her way to successfully *having* ϕ -ed, but that is precisely what is ruled out by Thompson's picture. Beyond that, if we go Thompson's way there is also nothing an agent can do with her knowledge of what she is doing that she can't do with her knowledge of what she intends. After all, if what the agent is doing and what she intends can come apart, then knowledge of

11. cf Setiya (2011).

the former can combine with knowledge of the latter to guide action: cueing steps, enabling course corrections when things go awry. But if they march in lockstep, this is impossible. If we go Thompson’s way, practical knowledge is just a byproduct, a spandrel.

It could of course be true that practical knowledge is a spandrel. But this would, I think, be surprising news to Anscombe, who made it central to her account—which is to say that I don’t think Anscombe herself would be delighted at this rescue.¹² And while we needn’t take fidelity to Anscombe as our mission, it seems to me that the particular way that Thompson’s picture violates the spirit of her view really is worrying, in that it seems to rob us of a resource that we had in hand before we began. We started by thinking practical knowledge might be rare, but valuable when had (and theoretically useful); Thompson leaves us with truckloads of the stuff, along with the sad news that it is worthless.

I want to say a bit more about this objection, but first I should get Setiya’s very different picture on the table, because it will turn out that my objections to the two views are symmetrical.

Where Thompson saves the knowledge condition by weakening the content that must be known, Setiya instead weakens the attitude: Roughly, he thinks that if an agent is ϕ -ing intentionally, she is *exercising a capacity* to know what she is doing in Anscombe’s special way, although on a particular occasion she may merely believe, or she may fail to believe yet be more confident than she otherwise would be, or have some other, weaker but still cognitive attitude in the neighborhood.¹³

12. In fact Anscombe said (in her §29) that “to say that I really ‘do’ in the intentional sense whatever I think I am doing” is a “false avenue of escape.”

13. This way of putting the point—in terms of the exercise of a capacity—is found in, e.g., Setiya 2016 (9–12); see also see also his 2011, 2012. For the original version and various refinements that lead to this point, see his 2007, 2008, 2009, etc.

So, unlike Thompson, Setiya does take the counterexamples to be counterexamples. But he invites us to notice that there are a range of possibilities compatible with the exercise of this capacity to know. For starters, the carbon copier and the other problem cases presented involve quite *complex* actions—and Setiya (2007) suggests that while the carbon copy maker (for example) does not believe he is making ten good copies (but only that he is trying to do so), he believes and even knows that he is doing some of the more basic constituent actions: things like *writing heavily*, and *pressing down hard*. Therefore, Setiya (2007, 26) initially proposes a condition that captures this idea:

Setiya’s BELIEF: When you are acting intentionally, there must be something you are doing intentionally, not merely trying to do, in the belief that you are doing it.

The ‘not merely trying to do’ condition means that the carbon copier’s belief (even knowledge) that he is trying to make ten copies is not enough for his making ten copies to count as intentional; he must be doing, and believe he is doing, the *writing heavily* and the *pressing down hard*.

This proposal is not far from something Davidson concedes: as Setiya (2011, 25) notes, Davidson (2001, 50) acknowledges that “Action does require that what the agent does is intentional under *some* description, and this in turn requires, I think, that what the agent does is known to him under some description” (emphasis added). Unfortunately, it is not yet a solution. Setiya himself (2011, 172) offers a Davidson-style counterexample:

Suppose that, as I recover from paralysis, my hopes are modest.
I think that I might be able to clench my fist, without being sure.
When I try to do so, I succeed: I clench my fist intentionally.

Here we have a lack of belief coupled with successful intentional action, as

in the carbon copier case, but the action in question is a basic action.¹⁴ There is no constituent part available to satisfy BELIEF. And Setiya offers other versions: for example, I might be not recovering from paralysis, but merely anæsthetized, with my hand behind my back. To handle this kind of case, Setiya (2008, 2009, 2012) revises his principle, taking advantage of the machinery of partial belief:

Setiya's BELIEF*: If you are ϕ -ing intentionally, you believe that you are ϕ -ing, or you are more confident of this than you would otherwise be, or else you are ϕ -ing by doing other things for which that condition holds.

The thought is that even in the paralysis or anæsthesia cases, you will at least be, in some perhaps negligible way, more confident that you are clenching your fist as you (in fact, it turns out) do so intentionally than you would otherwise be. This is a much weaker requirement than Anscombe's thought, but it preserves the cognitive flavor of intention while doing a much better job on the extensional questions.

Even this very qualified principle is controversial. For one thing, what about cases where what I intend to do is breathe normally? It seems plausible that I should be less confident that I am breathing normally when I am trying to do so intentionally than when I let my autonomic system take care of it.¹⁵ Because specifying any particular necessary attitude faces challenges like this, Setiya (2012, 18) has come to prefer the 'exercise of a capacity talk.' As he says:

The exercise of this capacity [for practical knowledge] interacts with evidence of other kinds and may affect one's degree of confidence without sufficing for knowledge or belief. That is what

14. Though see Thompson (2008) against the possibility of basic action.

15. The case is Setiya's own, from his (2009) reply to Paul; see p. 130.

happens in the case of recent paralysis, where I am clenching my fist intentionally but only partly believe that I am doing so because I am not sure that I've recovered. Such examples do not refute the Anscombean conception of the will, any more than examples of perceptual uncertainty refute the idea of perception as a source of knowledge.

I am sympathetic to the idea that the will is a capacity for knowledge; hence the title of this paper. But I want to suggest that it is a significant cost to weaken the attitude required in the content of the principle in search of a workable necessary condition, even if we suppose that Setiya's principle is descriptively and extensionally accurate. In §§1.4–1.6, I will offer and argue for an alternative principle that keeps the focus on the connection between intentional action and *knowledge*, instead of weaker attitudes, while still avoiding the counterexamples.

One way to understand what I am up to in this paper is to say that I am drawing out an undeveloped implication of Setiya's way of thinking: I am suggesting that this capacity for knowledge is not idle, not a mere interesting correlate; rather, it lies at the heart of agency. My principle focuses on how when an intentional action fails to realize the agent's capacity for practical knowledge, that action is therefore defective *as an exercise of agency*. And this thought is what we need to see my objections to Thompson and Setiya as parallel.

On Thompson's view, it's easy for me to know what I'm doing in part because it doesn't matter whether I will succeed in that doing: he makes agential success unnecessary for epistemic success. On Setiya's view, it's the other way around. He counts as agential successes cases where there is mere belief, or even just a slight uptick in confidence: everything can be going just fine for me as a doer, even when my work as a knower is not so impressive.

So for both Thompson and Setiya, success in knowing and success in doing vary independently, and one can be doing well, even maximally well, in one while things are going horrendously with the other. What I want to suggest

is that this fails to capture the key connection between knowing and doing. As I said regarding Thompson, on his picture, practical knowledge doesn't seem to be of any use to the agent in her acting. And on Setiya's picture, it's not clear if, or why, the carbon copier is doing less well as an agent in his doubtful making of ten copies than when he makes only two in full confidence. I think it is right (and more in the spirit of Anscombe's thought, for what it's worth) to think that we should not hold either side fixed while the other varies. And so in what follows, I will argue for a principle that ties together agential and epistemic success, in such a way that both are graded and they scale together. The essential thought will be that epistemic success and agential success vary together because—it turns out—practical knowledge is required for agential control, and control is required for agential success. I can't be in the good case unless I succeed in having ϕ -ed *and* I know I'm ϕ -ing, because what we want from an exercise of agency can't be had without practical knowledge.

1.4 MY PROPOSAL: MIND CONTROL

My proposal takes the following thought as its starting point:

MIND CONTROL: Whatever else it may be, an intentional action is an exercise of agency such that something the agent has in mind is supposed not only to 'match up' with something that happens in the world, but to guide and secure that happening.

MIND CONTROL is not a definition, but something like it has to be right if we are to be able to distinguish intentional actions from nearby phenomena. If we omit the requirement that the thing the agent has in mind *guide and secure* the action, requiring only that mind and world match, we do not distinguish intentional actions from (e.g.) mere wishes, or idle predictions, coming true. If we keep *guide* but omit *secure*, then at best we do not distinguish success achieved via agential control from merely accidental, or lucky, success; and at worst we do not even distinguish successful from merely attempted actions, or capture the important way in which a failed

attempt is defective. If we keep *secure* but omit *guide*, we do not distinguish proper intentional actions from cases of deviant causation.

MIND CONTROL suggests a rough, functional understanding of intention:

Intention: the mental state whose role is to guide and secure the agent's ϕ -ing, thereby (in the good case, when hooked up with the world in the right way) constituting agential control.

So the key feature that distinguishes intentional action is this connection to control.¹⁶

1.5 CONTROL AS PRACTICAL KNOWLEDGE

The truth of MIND CONTROL entails some interesting things (some of whose truth and/or interest will require a bit of drawing out below). It's worth stating them here, together, to get them on the table as a group before discussing them.

Gradability: Control comes by degrees—and once you have an aim that comes by degrees, you have an *internal standard of assessment*.

Content: The guidance requirement determines the kind of content intentions must have; specifically, an intention must capture not only an outcome, but also the agent's means and reasons, if any. Guidance means intentions describe not just what, but how, and why.

Hookup: The securing requirement is about the relationship between the intention and what happens in the world; it is an *anti-luck* condition.

Given *Gradability*, we must (for starters) modify Anscombe's claim—which provides a necessary condition—to make it a normative condition. But once we do, since the aim of control is constitutive of intentional action (as we saw in §1.4), we'll have a standard that tells us how an intentional action can be

16. Since I will not, in what follows, be discussing these other phenomena any further, I will, for brevity, often say simply 'actions' when I mean 'intentional actions.'

better or worse *qua intentional action*. This is not to say that this standard gives us the only way that an action can be defective—there may well be independent standards of ethics, or prudence, or æsthetics, for example—but it is a standard, and (the) one that is internal to intentional action as such.

Why think *Content* is true? Roughly, the idea is that for guidance, you need three things.

1. A starting point: you have to know where you are. Your starting point is your circumstances; a subset of these, containing some especially salient features of your circumstances, is your reasons (if you have any)—the *why*.¹⁷
2. A destination: you have to know where you're trying to go. Your destination is the outcome—*what* you intend to do, or (perhaps better) to have done.
3. A route: you have to have a path from the starting point to the destination. Your route is your means (if you have any); this is the *how*.¹⁸

I take it to be obvious that the outcome and means must be included for guidance; after all, what is guidance, if not a laying out of the steps to take to move from point A to point B? But, one might wonder, why think that *reasons* (when present) must be part of the content of an intention, as I claim? Why not think one merely intends *to φ*, rather than *to φ*

17. I will leave open the possibility of intentional action for no particular reason, as Anscombe does. (Though if such actions exist I think, with Anscombe, that they must be marginal cases; see her §§20–22.) The point here is, first, that the agent must have a sense of her circumstances—to take oneself to have no reason is in part to take oneself to be in certain circumstances—and, second, that if the agent has reasons they must be captured in this way, and that the lack of any reason must likewise be captured.

18. Again, I do not wish to rule out the possibility of basic action, action for which no means are required. The point here is analogous to the discussion of reasons above: if there is a means that is to be taken, it must be captured; if there is no means to be taken, that feature must likewise be captured.

for reason(s) *R*? The answer is slightly complicated because of the many kinds of reasons there are—recall the different types of answer to the ‘why?’ question discussed in §1.1—but in a nutshell, the answer is that reasons *constrain* both what counts as the intended destination and what routes are preferable, acceptable, and even viable.¹⁹

Compare two reasons why I might intend to drink a martini:

Election Martini: I intend to drink a martini *because of the election*.

007 Martini: I intend to drink a martini *to look like James Bond*.

Election Martini and *007 Martini* differ in what counts as doing what I intend: in *Election Martini*, I will not have succeeded unless I drink the whole thing; in *007 Martini*, I may well need only to take a sip between turns at the baccarat table. That is, the difference in reasons makes for a difference in goals, even though they can both be loosely characterized as ‘drinking a martini.’

Election Martini and *007 Martini* also differ in what routes are preferable and even viable paths to the destination: in *Election Martini*, my intention will recommend starting by obtaining a *very large* martini, the bigger the better, which I will be content to drink from a plastic cup while wearing sweatpants. In *007 Martini*, I will not particularly care about the size of the drink, but I will care a lot about its arriving in a proper glass; I’ll prefer to be wearing a very well-tailored dinner jacket; and I will make sure I say “shaken, not stirred” while ordering. If my reason for drinking a martini is to be like James Bond, then the means that would work just fine in *Election Martini*—plastic cup, sweatpants—will in fact not be a viable means to my end. So reasons, when present, must be captured by the intention if the intention is to mark off a destination and delineate proper means, as guidance requires.

19. Cf Fara 2013 for some analogous thoughts about desire.

Do the martini cases really mean that one's reasons must always be part of one's intention? I have shown that *there exist cases* in which the reasons matter in the way I claim; the existential claim does not obviously establish the universal one. But in fact, I think here it does.

From perspective of the intending agent, since reasons can *sometimes* affect what outcomes or what means are preferable, acceptable, or viable, you've always got to check: you can't know ahead of time in a particular case that reasons don't matter. Suppose I intend to wear my striped shirt to campus today. I might intend to do so because it's my lucky shirt (call this reason *Lucky*), or I might intend to do so because the checked one is in the laundry (call this reason *Laundry*). Both those reasons, it turns out, will have the same guidance effects: I should go to the closet, get out the striped shirt, put it on (along with whatever else I'm going to wear), and leave for campus. *Lucky* and *Laundry* seem equivalent with regard to guidance. But I might have intended to wear my striped shirt because I've already got it on and I don't feel like changing (call this reason *Lazy*)—and *Lazy* does not guide me in the same way: the whole point, in the *Lazy* case, is that I'm not going to go to the closet before I leave for campus. I have to include the reason to know, if I intend to wear my striped shirt to campus today, what that means as far as whether or not I need to go to the closet. Because reasons can offer different verdicts about guidance, it doesn't matter that sometimes they will in fact agree. One could intend to ϕ for all kinds of reasons; some of those reasons will very likely disagree about guidance. So you have to 'check.'²⁰ Different reasons make for different intentions, because reasons are part of the content of an intention.

The shirt case is still a complex action: do I really need to check if I'm performing a simpler one? What about a basic action, if there are any?

20. I am not suggesting that agents do, or must, *consciously* check; I mean only that the way reasons interact with the other constituents of an intention make it the case that the reasons must be incorporated into the intention.

Suppose I intend to blink, or clench my fist. Even those can be done in different ways—quickly or slowly, forcefully or gently, etc—and the reason I am blinking or clenching could bear on how I proceed. And even if there is an action that can only be performed in one way, the question of when to do it, or whether to do it at all, can easily be affected by the reason for which one intends to do it. Recall that drinking a martini out of a red plastic cup while wearing sweatpants will not count as doing what I intend if my reason is *to look like James Bond*, so my reason would thus recommend that I either forego drinking the martini, or wait until it can be done in the right way, if the only way I can drink a martini (now) is in the *Election Martini* way. The same can be true for a basic action.

So I would suggest that the right way to think about intentions is as having ‘slots’ for the what, the how, and the why, even if sometimes we want to say there will be no means, or no reason. Roughly, I intend to ϕ by means m because r . Notice that thinking this way makes it clear why there are the several kinds of answer to the ‘why?’ question described in §1.1: If you ask why I am ϕ -ing, I may reply with something about my means, something about my reasons, or something about how actions and sub-actions fit together. And in that sense, items that fit in any of the slots can be a reason, because answers to ‘why?’ questions are reasons for acting, as Anscombe says.

Therefore, I take it, an intention must include an agent’s reasons, and guidance explains why. Of course, it is also independently plausible that one does not merely intend to ϕ , but rather intends to ϕ *because* r . Just consider how many domains outside action theory—ethics, epistemology, decision theory—take it to be obvious that there is very often a crucial difference between ϕ -ing because p and ϕ -ing because q .

What about *Hookup*? *Hookup* does two things. First, it notes that if an intention must (in the good case) *secure* an outcome, then MIND CONTROL is in part an anti-luck condition. Intentional actions are defective *qua* intentional actions to the extent that they are achieved through luck (as opposed

to agency). Second, what makes this particularly interesting is that the kind of luck that must be avoided is *epistemic*; it is the kind ruled out by *knowledge*. We can see all this by thinking about skill.

To begin, notice that skillfulness is a kind of assessment specific to intentional actions; we do not evaluate other phenomena (such as those discussed in §1.4) in terms of skillfulness.²¹ Skillful actions are thus better *qua* actions than less- or non-skillful actions.

What's more, skill amounts to a reduced reliance on luck; the success of a (more) skillful action is less lucky than the success of a (less- or) non-skillful action. Serena Williams is extremely skilled at serving a tennis ball; I am woefully unskilled at serving a tennis ball. For any given pair of instances of serving a tennis ball, if Serena's serve goes in, or is an ace, it is not (or not much) a matter of luck; if my serve goes in, or is an ace, it is a matter of luck, to an embarrassing degree.

Of course, we might think even Serena needs *some* luck; she needs the world to refrain from providing a rogue tornado at the moment she serves, for example. But any luck that Serena requires, any tennis player requires; it is plausible that she requires less luck than any other player, perhaps in the history of the game—and I require far more luck than Serena (or than an even halfway decent player). If my serve is to go in, I will need luck in the timing of my toss, the angle of my racquet head, the amount of force I exert, and so on and so on and so on. Every instance of a successful serve by Serena will be much less lucky than any instance of a successful serve by me, and in particular, systematic ways.

As this enumeration of the luck I need for a successful serve suggests, the reduction in luck involved in skillful action is brought about, for starters, by

21. Of course, *agents* are also assessed in terms of skillfulness. But I set that aside for now; it is no objection to what I say here, in fact if anything it speaks in favor of my point: skillful agents are, roughly, just those creatures prone to performing skillful actions.

knowledge of means (where we might gloss ‘means’ as ‘what to do when’).²² When Serena executes one of her skillful serves, the serve manifests her knowledge of how to hold her racquet, when to release the ball and when to swing her arm, how hard to swing, etc.²³ If Serena merely believed those things about her means, even truly believed them, that would be inadequate for the particular case of serving to be skillful, because then the success of the action would be lucky, and lucky success is not skillful success.²⁴ And if Serena were to succeed in serving an ace during the rogue tornado, that particular ace would not have been skillful, it would have been lucky: Serena does not in fact know how to modulate the force, angle, etc of her serve to compensate for the tornado (no human could; our perception doesn’t provide enough information about the states in the tornado, and the dynamics are too complex to calculate in time anyway), and given that she did not know the tornado was coming (we can stipulate that it was a rogue tornado that materialized with no warning after the instant when she hit the ball), even if she had known how to serve in a rogue tornado, she would not have served

22. The ‘what to do when’ phrasing is owed to Pavese 2016, which provides a very helpful overview of the recent debate about the relationship between knowledge and skill. Most of what I say here aligns with the view presented in Stanley & Williamson 2016, which explains skill in terms of knowledge, though it’s important to note that they focus on skillful agents, whereas I am primarily interested in skillful actions. Perhaps surprisingly, what I say here is also in many ways compatible with a rival view, the virtue-theoretic epistemology on which knowledge is defined in terms of skill (as in, e.g., Zagzebski 1996, Sosa 2007, Greco 2010). I suspect that the picture I am developing in this paper provides resources that bear on this debate, but I cannot pursue the question here.

23. Cf Stanley & Williamson 2016 on the many kinds of knowledge-*wh* that are involved in skillfulness.

24. If it’s hard to see this, because it’s hard to imagine Serena having merely true belief, try another example: Suppose I believed in 2015 that the way to get rich was to bet my life savings on the Chicago Cubs’ winning the 2016 World Series, because I noticed that my cat had a spot that looked kind of like the Cubs’ logo. If I had bet my life savings on the Cubs’ win, I would have gotten rich. But no one would say that it was a skillful instance of getting rich; it was too fluky.

in the appropriate way in this instance.

But knowledge that some means is a way to ϕ in general is not enough; to act skillfully, an agent must know a means that is a way for *her* to ϕ in *her* circumstances. Otherwise, it will be lucky that the chosen means, which works in general, works in this case. So in the good case, the agent's intended means is one that is sufficient for success by that particular agent, here and now. And the agent must *know*, in the good case, that her intended means is sufficient for success by her here and now; otherwise it will be lucky that she intends (chooses) the sufficient means rather than some other one. And that means what is needed is in fact *knowledge of ability*, knowledge that the agent's intention is sufficient for success here and now.²⁵ To see this, let's set aside the tennis examples and consider a different kind of case.

Suppose there is a massive, incredibly complex, but perfectly constructed domino chain before me. In fact, it is the longest domino chain ever created, and it is full of exciting and quirky and original components, culminating in a 100,000-domino portrait of Anscombe.²⁶ I am the greatest domino chain expert in the world; I built this wonder; and all that is left to do is to flick the first domino to start the chain—the absolute easiest part of the whole undertaking.

But today, it happens, I have recently recovered from paralysis after a particularly severe domino mishap, and I have not quite recovered my normal level of motor acuity in my flicking finger. I know the chain will do all the wonderful things it's supposed to if I just don't mess it up at that first step—that is my expertise, and I built the chain, we can suppose, before

25. Cf Setiya 2008 on the roles and importance of knowledge how and knowledge of ability in intentional action.

26. According to the Guinness Book of World Records, as of this writing the most dominoes toppled by an individual is 321,197, a record set by Liu Yang in 2011. It took more than a month to set up the chain of 321,200 dominoes, 3 of which did not fall. Guinness is silent on whether Anscombe was depicted.

my injury, so there is no worry that something is out of whack there. But there is a way I can fail to bring off my task: if I tap too hard, the first domino will jolt its successor in a way that will ruin the chain. If I tap too lightly, it will be moved out of place (and suppose I do not currently have the dexterity to put it back just so). So I only have one shot, and I know the only thing that can go wrong is if I fail to flick the starter domino correctly. All I need is to perform the basic action of flicking the domino with an amount of force in the acceptable range, but I don't know if I can do this, and so I don't know if I can successfully achieve the run. And so, I want to suggest, if I do succeed, my success will be lucky—and in a way that means it is not skillful, even though under normal circumstances it would be, and even though we might wish to say that I, as an agent, still have skill in this domain—because I do not have enough control of the basic action of finger-flicking.

What the domino case brings out is that skillful actions are *actions from knowledge*, and in the good case the agent has all the knowledge required to make it the case that if she intends to ϕ , she will ϕ : what is wanted is knowledge of ability, which rests on knowledge of circumstances plus knowledge of means (if any).

Let's put all these pieces together. (Into something that will be, indeed, rather Anscombe-shaped.)

- The requirements for guidance (in *Content*) show that an agent's intention must represent what she is doing, including (if applicable) her means and her reason(s).
- The requirements for securing (in *Hookup*) show that this representation must be accurate, and not as a matter of luck.
- Therefore, in the good case, because the agent must have knowledge of her ability to ϕ via her intended means, once she begins to ϕ in light of, or from, that knowledge, she has *knowledge in intention* of what she is doing. (I'll say more about why below.) Which is to say that to aim at agential control in the way we are considering is to aim

at Anscombean practical knowledge. Thus intention is the practical analog of belief.

Those who go in for a picture of intention along the lines Bratman (1987, 1991, 2009, etc) advocates, on which intentions are distinctively practical attitudes that involve planning and commitment and don't have this cognitive flavor, won't like the conclusion that MIND CONTROL entails an aim of knowledge in intention. Even if I'm right that control requires knowledge of ability, they may ask: why couldn't it be that the intention is something Bratmanian, and thus non-cognitive?

Recall that MIND CONTROL requires that *something* in the mind guide and secure what happens. In this section, I have argued that that job is only accomplished when the agent has (and acts from) knowledge of ability. If intentions are at least partly cognitive, as I claim, and so the practical analog of belief, it makes sense that they would aim at knowledge in this way. On the other hand, if intentions are non-cognitive, there is no explanation for why their constitutive aim (of control, which is practical knowledge) turns out to be epistemic.

To put it another way: in §1.4 I said that MIND CONTROL suggests a functional account of intention, on which intentions are the mental states tasked with the guiding-and-securing job. On that view, the aim of knowledge I derive is internal to intention. The Bratmanian will have to argue that intentions are non-cognitive, but nonetheless constitutively come with two curious add-ons: first, an aim that is epistemic, and therefore, second, some other accompanying mental state to do that epistemic work (in the good case), since there is nothing in a Bratmanian intention that is fit for purpose.

I have said a lot in this section, but we can now sum up. MIND CONTROL distinguishes intentional action from nearby phenomena by identifying a normative condition specific to intentional action; identifying agential control as that condition; and identifying agential control with knowledge in intention, or practical knowledge. Thus agential control, i.e. practical knowledge, is

the (or at least *a*) *constitutive aim* of intentional action. MIND CONTROL, along with its corollaries *Gradability*, *Content*, and *Hookup*, yields a principle that's in the spirit of Anscombe but more plausible (and not vulnerable to Davidsonian counterexamples, as we'll see):

NEED TO KNOW: When you act intentionally, you aim at control, understood as practical knowledge; when (and to the degree that) you fail to have such control/knowledge, your action is *defective*, though it can still be intentional.

At this point some people will think, *hang on, knowledge isn't degressed!* But there are a couple of things to say here. First, one way an agent can have more or less knowledge of what she is doing, and so more or less control, has to do with knowing more about her circumstances, means, etc. She can know these things in more or less detail, which is a matter of knowing more or fewer particular facts.

Compare me to God. Her intentions can be perfectly precise, which is to say she has total control, because she knows everything. For me there are lots of things about my circumstances I don't know, which means even if I perfectly knew how all possible means worked, there would still be some luck (because I wouldn't know I'd chosen the right means for my circumstances), and of course in fact there is a lot I don't know about how my means work, so it is (at least ordinarily) lucky to some degree when they do. Those things that I don't know add up to less-than-perfect control.

Second, we could also think about this question in terms of degrees of confidence: if an agent has less knowledge about her means, she is entitled to less confidence that she is ϕ -ing. (This way of thinking will be especially appealing to those who emphasize the importance of basic action, who are likely to resist building everything into the content.) I am inclined to think that in some cases, thinking in terms of levels of detail will be more illuminating, while in others we'll want to think in terms of degrees of confidence—but in any case, these two possibilities make it clear that knowledge of what one is doing comes by degrees.

1.6 REVISITING DAVIDSON'S CARBON COPIER CASE

Let's revisit Davidson's carbon copier case in light of our discussion thus far. To fill it in a bit:

Abraham the Carbon Copier: Someone, call him Abraham, wants to make ten carbon copies of a page of text, all at once. (He has a lot on his plate and doesn't have time to write the thing out eleven times.) So he's built a stack of papers—eleven sheets of regular paper, interleaved with ten sheets of carbon paper—and selected his sturdiest pen. He begins to write, pressing down very hard: “Four score and seven years ago, our fathers brought forth on this continent a new nation...” Given that the odds of Abraham's succeeding in making ten copies all at once are very low, he doesn't believe he is doing so. Nonetheless, if he manages it, he will have done so intentionally: we would certainly not say that someone who built the stack of papers, selected a special pen, pressed down hard for the entire duration of the episode, etc, had made all those copies by accident.²⁷

Abraham is a counterexample to Anscombe's Thought: he does not know, because he does not even believe, that he is making ten carbon copies, even if in fact that is what he is doing, and intentionally so. But the case is no problem for NEED TO KNOW: according to NEED TO KNOW, Abraham's non-believing copying is intentional, but defective. And this seems right: it's why Thompson urged us to recognize how unlikely it is that an actual agent would attempt the whole shebang rather than checking as she went, as discussed in §1.2. Agents are very often sensitive, *responsive*, to defects like this.

But we can say more. Not only is Abraham not a counterexample to NEED TO KNOW, he is in fact predicted by it. After all, Abraham's skepticism

27. Abraham and the particular text here do not figure in Davidson's case, which is more schematic, and the five known handwritten copies of Lincoln's address vary in the details of their wording, and so were not in fact made with carbon paper. But carbon paper was invented in 1801, so the Great Emancipator *could have* decided to make some duplicates of his speech for posterity.

about his own action is a product of several gaps in his knowledge regarding his means, gaps we should expect to find among non-omniscient, non-omnipotent agents.

- *Could one do this, in principle?* Abraham doesn't know if there is any amount (or range) of pressure he could exert that would be enough for his stroke to reach all the way to the bottom sheet without tearing the top sheet(s).
- *How hard would one have to press?* Abraham doesn't know, if such an amount (or range) of pressure exists, what amount (or range) it is.
- *Do I have the strength?* Abraham doesn't know, if such an amount (or range) of pressure exists, if he's strong enough to exert it.
- *Do I have the motor acuity?* Abraham doesn't know, if such an amount (or range) of pressure exists, if he has fine enough motor control to target and maintain it over the course of the writing of the page of text.

There are likely many more crucial things Abraham doesn't know (whether his pen nib is adequate in the requisite ways, for example), but I'll stop here for now. What we see is that these issues are precisely the kinds of problems that motivated the requirement of knowledge of ability in the argument for §1.5. The first two issues capture Abraham's lack of knowledge of whether, and how, it is possible to make ten carbon copies at once: does there exist a means to his end? The second two issues concern whether, assuming *someone* could make ten copies, it is possible for *Abraham* to do so, and do so now. So we have four types of potential problem. The task might just be impossible, so that nothing Abraham does would be a way of accomplishing the task. The task might be possible, but Abraham might select the wrong way of going about it (by guessing wrong about how much pressure to use), and so fail. The task might be possible, and Abraham might select the correct way of going about it, but lack the requisite physical strength, and so fail. Or it could be that the task is possible, and Abraham selects the correct way of going about it, and he has the requisite strength, but he is

unable to exert that amount of strength here, now, for the duration of the task, and so he fails.

I think it's worth considering in some detail just what kind of problem each of Abraham's knowledge deficits poses on a view informed by MIND CONTROL and accountable to NEED TO KNOW.

Let's start with the first one: Abraham's task may be impossible. Many theorists, including Davidson, have thought there is something wrong with intending to do something (one knows to be) impossible. Davidson (2001, 100) says that "There can be no such intention." Often, the problem is described as one of incoherence. How should we understand this thought?

The view we've been developing provides the resources to see what's wrong with intending to do something you know to be impossible: if ϕ -ing is impossible, then there is nothing you could have in mind that could (guide and) secure your ϕ -ing. The intention to ϕ is therefore inconsistent with knowing it's impossible to ϕ .²⁸

Of course, Abraham's plight is not one in which he *knows* it is impossible to make ten carbon copies at once; he would presumably think that was a better situation than the one he is in, in that he would have no need to try. Does MIND CONTROL help us to understand what has gone wrong when one does not know, but merely fears, that what one hopes to do is impossible?

Consider a case that does not have all four of the potential problems we identified for Abraham, but just this one. There are actual cases like this; Andrew Wiles apparently thought, in the mid-1980s, that if it was possible to prove (what we now know as) the modularity theorem, and therefore

28. Of course, there are different flavors of impossibility, but that poses no problem for what I have said. I err if I intend the logically impossible, the metaphysically impossible, or (it seems right to say) the presently/for me/actually impossible—that's what Abraham is worried about when he wonders if he has the requisite ability, we might think.

Fermat's last theorem, he could do it.²⁹ And many champions, from Serena Williams to the competitive eater Takeru Kobayashi, presumably often say to themselves, "If anyone can, I can." But it is not only those with extraordinary abilities who can entertain, and rightly, such thoughts. At the moment, I can't recall if there is any tea in the pantry, so I don't know if it's possible for me to make myself a cup (in the next ten minutes, without going to the store, etc). But if there is tea, then I know I can make it, just as many millions of others could, too, were any of them in my kitchen. Anscombe herself (1963, §32) mentions a case like this: someone who wants to buy tackle for catching sharks in Oxford. No special skill is required to pull that off, if it's possible at all; it's just doubtful that the stuff is to be had.

So let's think about Anscombe's case. She calls it an instance of an *error in judgment* to intend to buy shark tackle in Oxford, when one might have known such a feat is impossible. We've already seen that it would be a mistake to set out so intending if one in fact does know that shark tackle cannot be bought in Oxford, but what about the case where one is not sure?³⁰ Call our would-be shark catcher Liz. If Liz intends to buy her tackle in Oxford, and she knows (at least implicitly) that doing so would be incoherent if doing so is impossible, then in intending to buy her tackle in Oxford, Liz presupposes that it is possible to do so. In presupposing that it is possible to buy shark tackle in Oxford, Liz is *taking for granted* that it is possible to buy shark tackle in Oxford. But Liz's actual attitude toward the possibility of buying shark tackle in Oxford is not one that supports taking this for granted. And so her combination of attitudes—the intention and the degree of belief in the possibility of buying shark tackle in Oxford—is unstable.

29. Or so Simon Singh (1998) reports.

30. I confess, were I in Oxford, I would be tempted to attempt to buy shark tackle, just to see if an anti-Anscombean shopkeeper had started offering it.

But instability of this kind is incompatible with intention's role as a guide: a guide is supposed to be *fixed*. (Think of the rigidity of train tracks.) This instability is a defect. Another metaphor: consider the way one uses a ruler as a guide in drawing a straight line. If the ruler is not held steady, it can't serve as a proper guide (and the more unsteady it is, the worse it will be as a guide). It is thus less likely that the outcome will be as intended—that the line will be straight—and if the line is straight, this outcome will not have been secured by the use of the ruler in the way it would have been if the ruler were steady. We might want to say it will be *lucky* if the line comes out straight. To cash out the metaphors: the presupposition and the actual attitude prescribe different courses of action. The presupposition that shark tackle is to be had in Oxford supports Liz's heading out into Oxford and visiting each and every store until she finds some.³¹ Liz's actual attitude, which we are stipulating for the moment is one of uncertainty as to whether shark tackle is to be had, supports very different behavior. Indeed, it may make far more sense for her to try Amazon instead.

The same kind of instability will afflict someone who has Abraham's third problem: the "I know it's possible, but I don't know if it's possible for me" case. The intention is supposed to guide *my* action; if Abraham knows it would take x lbs of force to press through to the bottom sheet on his pile, but he knows his maximum force is $(x - 1)$ lbs, then there is nothing *he* could have in mind that would secure the outcome of ten copies made. This is the analog of known (total) impossibility. What about the case where the possibility-for-me is unknown? Again, there are plenty of real world cases: take the feats of Andrew Wiles and Serena Williams and Kobayashi: I know they are possible, because those people have accomplished them. But I am in serious doubt about my ability to prove Fermat's last theorem, or win even one game in a Grand Slam, or eat even a few hot dogs. But that doubt creates the same kind of instability in this case as in the shark tackle case.

31. Recall Williamson (2000, 62ff) on the dedicated burglar.

Intending to ϕ is presupposing not just that it is possible to ϕ , but that it is possible for me to ϕ . And the more doubt I have about my ability, the more unstable, and so defective, my combined attitude is.

What about Abraham's second problem, his lack of knowledge of which way or ways of proceeding are ways of making ten carbon copies, assuming that it is possible to do so? We can see that it is incoherent to intend to ϕ by doing something that is not a way of ϕ -ing. And as with things that are known to be impossible, most everyone will agree that an intention to ϕ by doing something that one knows will not lead to having ϕ -ed is not okay. So, again, what about the case where you just don't know which way is the way to ϕ ? There are plenty of cases like this: you don't know which numbers to choose in order to have a winning lottery ticket; you don't know which is the road to Larissa. In cases like these, you are unable to discriminate on the basis of choiceworthiness between some range of options—perhaps just a few, if we are talking about possible roads to Larissa; some 175 million if we are talking about the Powerball lottery. This inability to discriminate means that if you choose the right one, your doing so will have been merely lucky, in exactly the way we have seen to be incompatible with control.

What of Abraham's fourth problem? His uncertainty about whether, even if he has the brute strength required to make ten copies, he can summon precisely the required amount—no more, no less—and maintain it over the course of an entire page of writing, can be put in general terms as the “Even if it's possible for me to do this (in one sense/in principle), can I do this (*tout court*/now)?” Once again, we can think of cases where this is the primary issue: Serena knows she can serve an ace to win a match, but even she doesn't always know she can do it *now*.³² Once again, we can see how the known inability version of the case would be defective; it looks like

32. It's worth noting here that a case in which a recently paralyzed person intends to clench her fist but is (properly) not confident she can do so—as in Setiya (2007, 2008)—is some combination of the “is this possible for me?” and “is this possible for me now” types. That case is simpler, because the action is basic. But the issue is the same.

our earlier examples of impossibility of various kinds. But in the version where the agent merely doubts whether she can ϕ now—which I glossed, in Abraham’s case, as not knowing if he had enough motor control to target and sustain the requisite pressure—it’s easy to see that the problem persists, because this is the domino flicker case again.

The case of Abraham the carbon copier points to each step in the argument in §1.5 that skilled action requires knowledge of ability; NEED TO KNOW is the product of just the problems present here. Thus NEED TO KNOW, although Anscombean in spirit, is not vulnerable to the counterexamples that doomed Anscombe’s original claim. More than that, it helps us understand just what is going on, and going wrong, in the case.

1.7 CONCLUSION

Time to take stock. I have argued that we need MIND CONTROL to distinguish intentional actions from nearby phenomena, and that this means that intentional actions constitutively aim at control. I’ve argued further that that control is just practical knowledge, in Anscombe’s sense: the special knowledge an agent can have of what she is doing, how, and why. So NEED TO KNOW is true: an intentional action is defective, *qua* intentional action, when, and to the extent that, the agent fails to know what she is doing (and how and why). I’ve also said that a Bratmanian view of intention does not appear to be able to account for this fact, so the arguments of the paper provide new support for cognitivism about intention.

Linking knowledge and non-defective action as my view does leads naturally to questions in epistemology about whether knowledge can come and go depending on non-epistemic features like what is at stake, leading to pragmatic encroachment, or making knowledge somehow *shifty*. In another paper, I explore what my view adds to that debate. Relatedly, if knowledge is essential for action in the way my view claims, we might wonder what that means when there are injustices in our epistemic economy. I take that question up in further work as well.

Finally, recall that I started by talking about God and Wile E. Coyote, the ideal agent and an agent with a lot of terrible ideas. Because my view identifies a constitutive aim of intentional action that can be met by degrees, it sets a standard for ideal intentional action: perfect practical knowledge. We might wonder how this standard relates to other standards, such as those of practical rationality, or ethics. Are all three of God's perfections—omnipotence, omniscience, and omnibenevolence—bound up together? Does NEED TO KNOW somehow implicitly tell us all that is required for good action? And what are non-ideal agents like Wile E. and this author supposed to do, frail and fallible as we are? If you've got to know what you're doing, but that's really hard, is there any way to avoid the ACME portable hole? These are questions for another day, but the seeds of an answer are here.

REFERENCES

- Anscombe, G.E.M. (1963) *Intention*. Cambridge, MA: Harvard University Press.
- Bratman, M. (1987) *Intention, Plans, and Practical Reason*. Cambridge, MA: Harvard University Press.
- . (1991) “Cognitivism about practical reason (review of *Practical Reflection*, by J. David Velleman).” *Ethics* 102(1): 117–128.
- . (2009) “Intention, belief, practical, theoretical.” In *Spheres of Reason: New Essays in the Philosophy of Normativity*, ed. Robertson, S. Oxford: Oxford University Press.
- Davidson, D. (2001) *Essays on Actions and Events*. New York: Oxford University Press.
- Fara, D.G. (2013) “Specifying desires.” *Noûs* 47(2): 250–272.
- Gibbons, J. (2001) “Knowledge in action.” *Philosophy and Phenomenological Research* 62(3): 579–600.
- Greco, J. (2010) *Achieving Knowledge: A Virtue-Theoretic Account of Epistemic Normativity*. Cambridge: Cambridge University Press.
- Hampshire, S. (1959) *Thought and Action*. New York: Viking.
- Pavese, C. (2016) “Skill in epistemology I: skill and knowledge.” *Philosophy Compass* 11(11): 642–49.
- Setiya, K. (2007) *Reasons without Rationalism*. Princeton: Princeton University Press.
- . (2008) “Practical knowledge.” *Ethics* 118(3): 388–409.
- . (2009) “Practical knowledge revisited.” *Ethics* 120(1): 128–137
- . (2011) “Knowledge of intention.” In *Essays on Anscombe’s Intention*, ed. Ford, A., Hornsby, J., & Stoutland, F. Cambridge, MA: Harvard University Press.
- . (2012) “Knowing how.” *Proceedings of the Aristotelian Society* 112(3): 285–307.
- . (2016) *Practical Knowledge: Selected Essays*. Oxford: Oxford University Press.

- Singh, S. (1998) *Fermat's Enigma: The Epic Quest to Solve the World's Greatest Mathematical Problem*. New York: Anchor Books.
- Sosa, E. (2007) *A Virtue Epistemology: Apt Belief and Reflective Knowledge* (vol. 1). Oxford: Oxford University Press.
- Stanley, J. & Williamson, T. (2016) "Skill." *Noûs* 50(3): 1–14.
- Thompson, M. (2008) *Life and Action: Elementary Structures of Practice and Practical Thought*. Cambridge, MA: Harvard University Press.
- . (2011) "Anscombe's *Intention* and practical knowledge." In *Essays on Anscombe's Intention*, ed. Ford, A., Hornsby, J., & Stoutland, F. Cambridge, MA: Harvard University Press.
- Zagzebski, L. (1996) *Virtues of the Mind*. Cambridge: Cambridge University Press.
- Williamson, T. (2000) *Knowledge and its Limits*. Oxford: Oxford University Press.

2 KNOWING WHAT YOU'RE DOING WITH SHIFTY EPISTEMOLOGY

If you're interested in knowledge and action, and what they might have to do with each other, there are plenty of epistemologists who will have something to say. There are also a bunch of action theorists who'd be happy to bend your ear. Oddly, though, the knowledge people and the action people have not been talking to each other. That's a mistake. (Especially for the epistemologists, it turns out.)

Action theorists who think action is importantly connected to knowledge are inclined to agree with Anscombe that there's something special about knowing what you're doing. Epistemologists who think knowledge is importantly connected to action generally say the link makes knowledge *shifty*, prone to coming and going depending on practical matters like what's at stake. This paper puts these lines of thought in conversation for the first time.

The shifty epistemologists start with cases. If you and I are talking about where we were for a historic event like the release of Beyoncé's *Lemonade*, we of course have elaborate answers, loaded with detail: we know *exactly* where we were. That is, unless a demonic game show host appears and says we'll get a new toaster if we're right about where we were, but if we're wrong we'll be forced to engage in online debates with Macedonian fake news bots for

the rest of our lives—then maybe we don't know after all.¹ Cases like this demonstrate, according to the shifty theorists, that knowledge can come and go just because of changes in one's practical situation, even if one's evidence is held fixed.

But this is surprising: after all, just as it would seem like bad epistemic practice to suppose that the fact that you would very much like for something to be true bears on whether you know it—treating wishful thinking as epistemically legitimate—it would seem that the more general question of how practically significant it is for you that something be true should not affect whether you know it. Isn't knowledge supposed to be solid, sturdy, dependable? The shifty theorists say that's just not true; the cases show that knowledge is shifty—and the cases are ubiquitous.

Anscombe's picture also starts with an intuitive thought and ends up with a surprising one. Notice that if you ask me why I'm typing this paragraph to the rhythm of 'Formation,' and I say that I didn't know I was doing that, then it seems right to say I wasn't intentionally typing to the rhythm of 'Formation.' If you don't know you're doing something, you're not doing it intentionally, she says.

But if that's right, then doing something intentionally means knowing you're doing it. And there are counterexamples to this claim—lots of them. Still, like the shifty epistemologists who say the cases compel us to accept that knowledge is not as stalwart as we had supposed, Anscombe's soldier on: intentional action really is marked by its connection to knowledge; it is we who must be stalwart in the face of the counterexamples.

In what follows, I argue that if we revise Anscombe's thought a bit, we'll have a way of thinking that illuminates the puzzles of shifty epistemology.

1. If this penalty strikes you as a description of the status quo and so not (merely) a consequence of a wrong answer, please substitute something horrific but non-actual. If you can think of anything.

It turns out that there are some questionable assumptions and troublesome ambiguities in the shiftiness literature, and once we have my modified Anscombean picture in view, we can see better what's at work in the shiftiness discussion—and in a way that makes shiftiness itself look unmotivated. In other words, we can take seriously the thought that knowledge is important to action without committing to knowledge's being shifty; we can think that lack of knowledge tends to hinder action *and* that whether you know something is a matter of traditional epistemic factors like your evidence, not your practical circumstances, despite the appearances in the shifty theorists' favorite cases.

To those ends, §2.1 lays out two kinds of cases that are the stock in trade of epistemologists' discussions of stakes-sensitivity, interest-relativity, pragmatic encroachment—more generally, the phenomenon of *shiftiness*—and attempts to get clear on the general structure of shifty views.² §2.2 revisits Anscombe's idea and its difficulties, and provides an overview of my preferred modification of her central claim. In §2.3, I explain how my picture of what intentional action is like yields principles—*knowledge norms* that link non-defective action to knowledge of certain propositions—that appear to be remarkably similar to those proposed by shifty theorists, but that in fact have crucial differences both in what they tell us about the cases and in their explanatory power; §2.4 walks through the comparison via the paradigm cases. Finally, in §2.5 I take stock, and suggest that while I have not provided a knock-down argument against shiftiness, I have shown that the cases that are supposed to motivate such views do not, in fact, push us that way—because my view about the relationship between knowledge

2. 'Shiftiness' is defined in Fantl & McGrath (2012, 55) as the claim that "factors such as practical stakes and salience can matter to the truth-value of 'knowledge'-ascriptions." This is supposed to be a (statement of the) claim that contextualists and interest-relativists, or subject-sensitive invariantists, both endorse. In this paper, I will almost always discuss the interest-relativist/subject sensitive invariantist versions of the claim, leaving the translation to the contextualist idiom as an exercise for the reader for the sake of concision. But the debunking arguments reach across the divide.

and intentional action, supported by a ground-up, independent argument, enables an equally satisfying explanation of the cases in question. Thus my paper can be understood as debunking shifty theories.

In short, it is illuminating to bring the two conversations about knowledge and action together; action theory can shed light on epistemology, shifting the ground under shiftiness itself. Because the point of this paper is that Anscombe was right that you've got to know what you're doing—and that means that if what you're doing is being shifty, you've got to wonder: why would I want to do that?

2.1 TWO KINDS OF CASES

I have said that the debate about shiftiness in epistemology centers on intuitions about cases; in fact, two main kinds of cases feature in the literature. I'll call them *criticism* cases and *encroachment* cases.

Criticism: An agent is criticizable when she acts on (the basis of) p , or takes p as her reason for acting, without knowing p .³ For example:

- *Restaurant:* Hannah and Sarah are on their way to a restaurant, where they have a reservation that they will lose if they are late. Hannah has a hunch that turning left will get them there, and so she leads them that way. It eventually becomes clear that this was the wrong way to go, and Sarah (naturally, we are to think) criticizes Hannah: “You shouldn't have gone down this street, since you didn't know that the restaurant was here.”⁴

3. This kind of case is at the center of Hawthorne & Stanley (2008), and since then has become commonly used, with some variation, in most of this literature. *Restaurant* is adapted from the opening paragraphs of that paper, and is followed there by several other instances of the type.

4. It's worth noting that on the views built atop this case, Hannah would be criticizable even if it had turned out to be the right way. And one might worry that the fact that it's the wrong way is what does the work in the case. But since this is a standard case,

Encroachment: Two agents, with (what is stipulated to be) the same evidence for p , act on (the basis of) p . The agents' practical circumstances are different—specifically, for one agent the stakes are high, and for the other agent the stakes are low—and (we are to think) an action that may be fine for the agent in the low-stakes case would be improper for the agent in the high-stakes case.⁵

- *Bank (low stakes):* Hannah and her wife Sarah are on their way home on Friday afternoon, and they plan to stop at the bank on the way, so they can deposit their paychecks. But as they drive up to the bank, they see that the lines are very long. There's no particular urgency to depositing their checks; they have plenty of money in their account. Hannah says, "I know the bank will be open tomorrow; I was there two Saturdays ago. So we can just deposit our checks in the morning when we go out to do the grocery shopping. Let's skip it tonight." Sarah nods.
- *Bank (high stakes):* Hannah and her wife Sarah are on their way home on Friday afternoon, and they plan to stop at the bank on the way, so they can deposit their paychecks. But as they drive up to the bank, they see that the lines are very long. There is some urgency to depositing their checks; they have very little money in their account at the moment, and their rent will clear on Monday. Hannah says, "I know the bank will be open tomorrow; I was there two Saturdays ago. So we can just deposit our checks in the morning when we go out to do the grocery shopping. Let's skip it tonight." Sarah frets, "But what if the bank has changed

I'll set the worry aside for now.

5. The pair of *Bank* cases here is adapted from the presentation in the introduction to Stanley (2005). Many encroachment cases are borrowed or adapted the contextualism literature; for example, the pair of train cases with which Fantl & McGrath open their (2002) is a very slight variation on Cohen's (1999) airport scenario.

its hours? Monday is the first of the month. Do you really know it'll be open?" Hannah replies, "Hmm. Maybe not."

2.1.1 THE CRITICISM CASES

Criticism cases like *Restaurant* are taken to be evidence that there is a *knowledge norm for action*: they seem to suggest that, at least, and roughly, you can properly act on the basis of p only if you know p .⁶

How are we to understand this claim? How is it supposed to be supported by *Restaurant*? Some unpacking is in order.

First: What does acting 'on the basis of p ' amount to? The 'basis' talk is from Fantl & McGrath (2009, 3), who say that when you know p , "you can put it to work... as a basis for belief [and] as a basis for action;" they also (see, e.g., their 2002, 77–8) talk about 'acting as if p ,' and explain the 'as if' in terms of rational preference—roughly, to act as if p is to take those actions it is rational to prefer if p .⁷ Hawthorne & Stanley (2008) talk instead about taking p as one's reason for acting. The shift from talking about properly acting to properly taking something as a reason is not insignificant, and I will have more to say about it below.

What is the sense of 'proper' action here? As the mention of rational preference above suggests, this is supposed to be a claim about practical rationality. Fantl & McGrath (2002, 2009, 2013, etc.) and Weatherson (e.g., 2012) explicitly have a decision-theoretic orientation; Weatherson is concerned with what can appropriately be written down on a decision table, for

6. I will often in what follows be abstracting from the details of and differences between various views in the neighborhood; for most of our purposes, we need only the broadly shared basic picture. But I will note some key divergences when relevant to my discussion.

7. These two versions do not seem equivalent; 'act as if p ' looks weaker. Fantl & McGrath's view has evolved over time; see the cited passages for the details. The difference should not matter for my purposes.

example. Stanley (2005), and Stanley & Hawthorne (2008), sometimes seem to have the same basic sense of propriety in mind, but sometimes seem to be after something different; again, I will return to this theme below. But I will set this thought aside for now, and understand the kind of practical rationality in question broadly enough to encompass both ways of thinking, if indeed they are distinct. As Hawthorne & Stanley (2008, 574n4) say in describing Fantl & McGrath's view,

The relevant notion of preferability, standard in decision theory, is one according to which A is preferable to B only if one welcomes the news that A to the news that B (“welcome the news” is a somewhat technical notion here, because in practice receiving the news that A typically involves receiving further information as well that might affect the preference ranking).

Although it seems natural to say that *Restaurant* and its cousins support only the claim that knowledge is necessary for proper action—after all, the criticism is roughly that since you did not know p , your acting as if (or on the basis of) p was improper—Fantl & McGrath (2002) suggest that the norm is in fact a biconditional. You can properly act on the basis of p *if and only if* you know p . The two directions come as a package, on their view, because of the way they understand acting as if p to be a matter of taking those actions it is rational to prefer conditional on p :

If you know that p , then it shouldn't be a problem to act as if p . If it is a problem to act as if p , you can explain why by saying that you don't know that p . Suppose you are faced with some decision—do A or do B—where which of these is better depends on whether p . You know that if p , A is the thing to do, but that if not- p , B is. To say in one breath, “I know that p ” and in the next breath, “But I'd better do B anyway, even though I know that A is the thing to do if p ” seems incoherent. If you really know that p , and you know that if p , A is the thing to do, then it's hard to see how you could fail to know that A is the thing

to do in fact. But then you ought to do A. (Fantl & McGrath 2002, 72)

The strangeness, perhaps incoherence, of saying that you know p while also saying that you ought to do something that is not the thing you ought to do if p is supposed to support the sufficiency claim. One might want a bit more argument, but by and large Fantl & McGrath think that the thought that, as they like to say, “if you know p , it shouldn’t be a problem to act as if p ,” is intuitive.

Hawthorne & Stanley (2008) likewise advocate for a biconditional norm, but do not offer much more by way of argument for the sufficiency direction. They do open the discussion of their preferred principle by noting that the necessity direction is what cases like *Restaurant* support. When it comes to the question of sufficiency, however, they are explicitly concerned only with suitably restricting the claim, so that their principle does not make knowing p sufficient to warrant actions for which the truth of p is irrelevant.⁸ In such cases, they say, “it would be odd to say that it is appropriate to treat the proposition that p as a reason for acting, even if one knows that p ” (578). Therefore, they continue,

To get a plausible sufficient condition, we need some restriction to choices for which the proposition that p is relevant. Let us say that a choice between options $x_1...x_n$ is p dependent iff the most preferable of $x_1...x_n$ conditional on the proposition that p is not the same as the most preferable of $x_1...x_n$ conditional on the proposition that not- p . (Hawthorne & Stanley 2008, 578)

In addition to addressing the irrelevance worry, one might think this is pointing toward the same connection to rational preference that is at work in Fantl & McGrath. (Otherwise, it looks as though the sufficiency direction is basically just assumed.) Again, we might have wanted a fair bit more; I’ll

8. We might add: or worse!

return to this question in §2.3 below, but for now notice that Hawthorne & Stanley's biconditional is not symmetrical. The necessity side, the left-to-right claim that if you can appropriately take p as your reason for action, then you know p , is unqualified; but the sufficiency side, the right-to-left direction, is restricted—because they think you can appropriately take p as your reason for action if you know it *and the truth of p is relevant, perhaps because it makes your action rational to prefer*.

2.1.2 THE ENCROACHMENT CASES

Encroachment cases like *Bank* can be read two ways. The first follows the use of such cases in the contextualism literature, and argues from the (expected) intuition that Hannah knows the bank will be open in the low-stakes case, and the (expected) intuition that Hannah does not know the bank will be open in the high-stakes case, even though their evidence is the same in both cases, to the claim that whether one knows depends in part on stakes. Call this the direct reading. The second reading proceeds from the (expected) intuition that it is rational to skip going to the bank in the low-stakes case, and the (expected) intuition that it is not rational to skip going to the bank in the high-stakes case, and adds a knowledge norm that is supposed to be plausible (perhaps because of criticism cases), to generate the conclusion that Hannah knows the bank will be open on Saturday when the stakes are low, but not when the stakes are high, and thus that whether one knows (that the bank will be open, e.g.) depends in part on stakes. Call this the indirect reading.

In encroachment cases like *Bank*, on the indirect reading, one is supposed to think that since (1) the agents have the same evidence in the high-stakes case and the low-stakes case, and (2) what it is intuitively appropriate for the agents to do in the high-stakes case is different from what it is intuitively appropriate for them to do in the low-stakes case (e.g., it is appropriate for Hannah and Sarah to skip the trip to the bank in the low-stakes version, but it would not be appropriate for them to skip it in the high-stakes version), and (3) you can properly act on the basis of p if and only if you know p ;

therefore, in a slogan, because practical matters affect what it is rational to do, “practical matters affect whether you know.”⁹ In other words, and more specifically: Suppose it is right that one can properly act on (the basis of) p if and only if one knows p . (Perhaps with a qualification requiring that p is relevant to the action in question, along the lines discussed by Hawthorne & Stanley.) And suppose that in *Bank* the relevant p is *The bank will be open on Saturday*. The thought is that if stakes alone make a difference to whether or not one can properly act on p , because by stipulation the stakes are the only difference between the cases, then—given the biconditional knowledge norm¹⁰—they must make a difference to whether you know p . Knowledge is thus *shifty*.¹¹

There are many questions one might have at this point; I have not here reproduced the full arguments for the various views in the vicinity. But I hope to have conveyed enough of how the line of thought goes to give us something to work with. And I want to flag a few things.

The key points seem to be:

- Criticism cases are seen to suggest a knowledge norm; the epistemologists who have liked that thought have generally read the action side of the relevant (bi-)conditionals in terms of practical rationality.
- Combining the knowledge norm with intuitions about the rationality of action in encroachment cases (as the indirect reading suggests) is supposed to show that, to put it very colloquially, it’s harder to know

9. “Practical matters affect whether you know” is in fact the title of a (2013) paper by Fantl & McGrath.

10. Notice that the sufficiency direction is required; the necessity side alone is not enough.

11. Again, Fantl & McGrath define shiftiness in their (2012, 55) as the claim that “factors such as practical stakes and salience can matter to the truth-value of ‘knowledge’-ascriptions,” where this is supposed to be a (statement of the) claim that contextualists and interest-relativists, or subject-sensitive invariantists, both endorse.

something (p) when more is at stake (for whether p).¹² Why? Because the norm says that if it's practically irrational for you to act on p (which high stakes may make the case), then you don't know p .

However, if the knowledge norm was not about practical rationality as these theorists suppose, then this link between having a knowledge norm and having knowledge be shifty might break. And that's where the picture of intentional action I've developed in other work comes in; I'll sketch it now.

2.2 INTENTIONAL ACTION'S AIM: PRACTICAL KNOWLEDGE

Anscombe said when you act intentionally, you *know what you're doing* in a special way, 'without observation' or inference.¹³ Remember how we began: if you ask me why I'm typing to the rhythm of 'Formation,' and I say that I didn't know that I was, then I'm not doing it intentionally.¹⁴ Add another thought: if instead I say, "Huh—I see that I am!" then I am still not intentionally *typing to the rhythm of 'Formation,'* even though I am probably intentionally typing. If I know what I'm doing, but only because, prompted, I *find myself* doing it, that's not the kind of knowledge

12. I noted earlier that many of the encroachment cases have their origins in the contextualist literature. And in fact there has been a substantial debate between contextualists and what I'll call interest-relativists about which is the right way to explain the apparent shiftiness present in these cases. Williamson (2005) helpfully crystallizes key differences between the two views; Stanley (2005) goes to great lengths to rebut the contextualist alternative. There are certainly differences, and not only because interest-relativists are concerned with knowledge while contextualists are concerned with 'knowledge.' But there are important similarities in the projects as well (cf Fantl & McGrath 2012): both camps are proposing shifty epistemologies. I take it that what I say here should apply equally, with appropriate translation, to contextualist accounts, but in the interest of concision I will not walk through how that goes.

13. See Anscombe 1963, e.g. 13–15, 49–50, for discussion of the non-observational, non-inferential nature of this knowledge.

14. See also Hampshire (1967, 95) on this point.

that distinguishes intentional action. Similarly, if my reply was “Hmm. I’m typing to the rhythm of the song Beyoncé sang at the Super Bowl. And the song she sang at the Super Bowl was ‘Formation.’ So I guess I am typing to the rhythm of ‘Formation,’” then it still seems like I wasn’t intentionally typing to the rhythm of ‘Formation,’ because I had to infer that that was what I was up to.

Anscombe took these data points and drew a natural (albeit strong) conclusion: intentional actions are knowledgeable (from the first point); the way this knowledge is had is, we might say, interestingly *direct*, as it is not derived from observation or inference (from the second and third points). This directness also means the agent’s special knowledge does not rest on sufficient prior evidence—where by ‘prior’ I mean ‘existing before the formation of an agent’s intention.’ This is true even though this special knowledge routinely relies *in part* on observation or inference; an agent will and must generally know many things about her circumstances, her abilities, etc., and she will know them in a variety of ways: the possession of such ordinary empirical knowledge is, for Anscombe, a necessary precondition of the special knowledge the possession of which is distinctive of intentional action. But the agent’s special, direct knowledge itself does not and cannot rest on that evidence alone. In other words, a candidate intention—a description of a possible action—will incorporate various pieces of the agent’s ordinary empirical knowledge. But when the agent converts that candidate intention from a possible intention into an actual one, the agent’s knowledge that *this* is what she is doing is not itself empirical.

If I’m intentionally typing to the rhythm of ‘Formation,’ then when you ask me why I’m doing that, I’ll say something like “Because that’s how you slay, paper-wise.” This is another face of the key Anscombean thought: when I act intentionally, I am able to answer a special kind of question; Anscombe (1963, 9) says, “A certain sense of the question ‘Why?’ is given application; the sense is of course that in which the answer, if positive, gives a reason for acting.” The knowledge that underpins my answer is Anscombean *practical knowledge*, an agent’s special knowledge of what she

is doing and why. The ‘what’ puts the agent in a position to accept the applicability of the question, and the ‘why’ is the answer. There is much more to say to get a full picture of Anscombean practical knowledge, and I will not take on that task here.¹⁵ What I want to emphasize is, first, the requirement of (a special kind of) knowledge—one that is bound up with the possibility of engaging with reasons—and, second, the way this special knowledge is, essentially, a non-accidental match between mind and world that is secured by the agent’s intention, her representation of what she is doing and why—because the intention not only describes something that happens, but also in an important sense causes that happening.^{16,17}

But there’s a problem: it is not true that all instances of intentional action involve the agent’s knowing what she is doing in this way: there are counterexamples, as Davidson and others have shown. For example, and perhaps most worryingly, I can intentionally ϕ without even believing that I am doing so, and thus without knowing it. Davidson’s (2001, 91–2) famous case involves a person making ten carbon copies at once, while in profound doubt that he is doing so. If you ask him why he’s making ten carbon copies at once, he’ll say “I don’t know that I am; I mean, I hope so, but what are the odds?” Setiya has noted that there are counterexamples of this kind even for basic actions like simple bodily movements—if I have recently had an episode of paralysis, I might intentionally clench my fist during an early stage of my recovery while being far skeptical about my ability to do so to

15. I will not address, for example, what makes this knowledge distinctively practical, though this practicality is crucial to the justification for the agent’s belief about what she is doing. See Setiya (2008), as well as Velleman (1989), for more on the epistemological question and its connection to the practicality of practical knowledge.

16. Just what kind of causation intention might involve is a difficult question, and I here set it aside.

17. For more detail about my preferred understanding of practical knowledge, see my (ms); my talk here of a ‘match between mind and world’ and my invoking intentions as mental states diverge from Anscombe’s behaviorist preferences, but are I think clarifying.

count as believing that I am doing it.¹⁸ So Anscombe's necessary condition does not hold.

2.2.1 SAVING THE ANSCOMBEAN THOUGHT

Other theorists sympathetic to Anscombe have suggested we should avoid the counterexamples by weakening one half or the other of the *knowing what you're doing* requirement—either weakening the attitude, requiring something less than knowledge, or weakening the doing-related content known.¹⁹ I suggest we should want more. We should want the full-strength attitude, and the full-strength action, and their success should vary together. What we should give up in deference to the counterexamples, I argue, is the necessity claim. An agent who acts intentionally has Anscombean practical knowledge of what she is doing *when things go well*; when (and to the extent that) she fails to know what she is doing, her action is *defective qua intentional action*.

In other words, knowing what you're doing is a normative condition on intentional action, not a necessary one. Anscombe was mistaken, but on the right track, in saying that when you act intentionally, you know what you're doing in a special way; sometimes you fail to have this knowledge, even as you act intentionally, as the counterexamples show. The right thought is that such failures occur in cases of intentional actions that are defective as such.

To state the principle I advocate:

NEED TO KNOW: When you act intentionally, you aim at *practical*

18. See, e.g., Setiya (2016, 41).

19. For the attitude-weakening approach, see Setiya (e.g. 2016, 9–12; 2012, 18). For the alternative, see Thompson (2008, 2011). It's worth noting that Anscombe herself seems not to have endorsed Thompson's strategy; see her 1963 (§29), where she said, "Another false avenue of escape is to say that I really do in the intentional sense whatever I think I am doing." I discuss these alternative views at length in my (ms).

knowledge, the agent's special knowledge of what she is doing, how, and why; when (and to the degree that) you fail to have such knowledge, your action is *defective*, though it can still be intentional.

This is not to say that an action is more or less intentional, depending on whether (or to what degree) the agent achieves practical knowledge. Rather, an intentional action, in virtue of being an intentional action, is *better or worse as such* depending on whether (or to what degree) the agent achieves practical knowledge.

The full argument for my view proceeds from a claim about how intentional actions are to be distinguished from nearby phenomena—ranging from wishes or idle predictions that come true, to merely failed attempts, to cases of deviant causation—and finds that it is precisely NEED TO KNOW that marks off the class of intentional actions. It is thus *constitutive* of intentional actions that they are defective when, and to the extent that, the agent fails to have practical knowledge of what she is doing (and how, and why). *What it is* to be an intentional action is to be subject to this condition.

So, again, what NEED TO KNOW captures is a *constitutive aim* of intentional action: what it is to be an intentional action is to aim at practical knowledge, which is to say that to be an intentional action is to be better or worse as such depending on whether (or to what degree) the agent in question has practical knowledge of what she is doing, and how, and why.

If the notion of a constitutive aim seems mysterious, it may help to think in terms of artifacts, because constitutive aims are closely connected to something like essential functions. So if I have a thermometer, it is better or worse, *qua* thermometer, depending on how well it measures the temperature. If a thermometer is very inaccurate, so that it does not in fact distinguish, say, 35° from 65°, it will be a very poor thermometer indeed, but it will still be a thermometer. The reason it is evaluable on the basis of how well it measures temperature is precisely *because* it is a thermometer; if it were instead a cup or a cookie, it would not be evaluable on the basis of

measuring temperature. There may be other ways we evaluate thermometers: by how easy they are to manufacture in quantity, or how beautiful or durable or inexpensive they are, or how significant the uses are to which they have been put. But what it is to be a thermometer is to be evaluable according to how well (or badly) one measures temperature, so measuring temperature is the constitutive aim of a thermometer.

The thought here is also precisely analogous to other constitutive aim talk in philosophical theorizing: for example, many people think it is constitutive of belief that it aims at truth. On this view, beliefs are defective when (and to the extent that) they fail to be true, and the existence and centrality of this mode of assessment *distinguishes* beliefs from other, similar phenomena, such as imaginings.²⁰ Being true is in some sense what beliefs are for, the thinking goes, just as measuring temperature is what thermometers are for.

On my view, intentional actions are distinguished by the aim of practical knowledge; intentions are the mental state whose function is to instantiate such knowledge. This function can be achieved to a greater or lesser degree—just as in the case of thermometers—so intentional actions likewise have a constitutive aim. It is the role of an intention to guide and secure an agent's action; intentions are the mental states that capture what the agent seeks to do, and how, and why, and then cause the agent to do that thing, in that way, for that reason. When, and to the extent that, things go well, so that an intention hooks up with the world as it should, then that intention constitutes practical knowledge, the agent's special knowledge of what she is doing, how, and why.

To sum up: While there was something appealing in Anscombe's idea that agents acting intentionally know what they are doing in a special way, her claim is subject to counterexamples, and so fails as a necessary condition.

20. The earliest statement I know of of the claim that belief aims at truth is found in Williams (1973); see also Railton (1994), and especially Velleman (2000).

But if we revise the claim, converting it to a *normative* condition as in NEED TO KNOW, we find we have a *constitutive aim*, and thus a standard that allows us both to distinguish intentional actions from nearby phenomena and to evaluate them as better or worse *qua* intentional actions.

2.2.2 A CLARIFICATION AND A CAVEAT

It's important to be clear that while I have argued that intentional actions have a constitutive aim, an aim by which we distinguish them from nearby phenomena and which provides an internal standard of evaluation, I have *not* argued for constitutivism about practical reason, the view—most famously advanced by Korsgaard (e.g., 2009) and Velleman (e.g., 2000, 2009)—that (all) the demands of practical reason are derived from the nature of agency. As with the thermometer that can be evaluated for cost, beauty, or historical interest in addition to its ability to measure temperature, so for all I have said we may evaluate intentional actions by any number of distinct, external standards of practical rationality, aesthetics, morality, and more, as well as by the internal standard of NEED TO KNOW. I have said nothing that requires that we get all our standards for action from the nature of agency; I have only suggested that we get *an interesting* standard for action from the nature of agency. In other words, for the purposes of this paper I am taking a position in the middle ground. To one side of me would be those who think there is no kind of normative standard constitutive of, or derivable from the nature of, agency; to the other, the constitutivists, who think agency is the source of all the other standards (whose otherness is thus merely apparent). It may be that this middle ground is unstable, and that the only live possibilities are all or nothing.²¹ It's worth noting that while I have not, as I keep saying, argued *for* a constitutivist view, neither have I ruled out the possibility of developing one from the materials here at hand. There is more to say on

21. cf Setiya (2013, 2014).

this subject, but that is a project for another day.²²

2.3 A DIFFERENT FLAVOR OF KNOWLEDGE NORM

The argument sketched above, leading up to NEED TO KNOW, provides a ground-up theoretical underpinning for a principle that makes good sense of both criticism and encroachment cases, but in a very different way. Here's what I mean.

NEED TO KNOW says intentional actions are defective when, and to the extent that, the agent fails to have practical knowledge of what she is doing (and how, and why). And that is already to say that intentional actions are subject to a knowledge norm: practical knowledge (of what one is doing, and how, and why) is necessary for non-defective action by the lights of NEED TO KNOW. One might wonder: is it also sufficient? Yes. The argument for NEED TO KNOW built up to the requirement of practical knowledge by thinking about what would be needed in the good case, for an agent to do well in acting intentionally as such. The thought was that intentional actions differ from similar phenomena in virtue of aiming at a particular kind of non-lucky match between what the agent has in mind and what happens in the world, and it turns out that that aim is achieved when the agent has practical knowledge, because then her intention is both an accurate description of what happens, and non-luckily so, because the intention is in an important (if tricky) sense the cause of what happens.²³

Given that, to say that practical knowledge is insufficient for non-defective action would be to say that the argument for NEED TO KNOW does not go through, not to say that practical knowledge is necessary but insufficient for non-defective action. To put it another way: the aim captured in NEED

22. I take the project up in my (forthcoming).

23. Again, I will not take up the many interesting and thorny questions about causation and intention.

TO KNOW is like the aim of getting out of the rain. I can have that aim and fulfill it only partly by ducking under a sparsely-leafed tree, so that I'm getting wet at a slower rate but definitely still getting wet; but once I step inside a building, with a proper roof and walls, so that I'm not getting wet anymore at all, then I've satisfied my aim of getting out of the rain, and nothing more is required—it wouldn't improve my performance relative to the aim of getting out of the rain to go deeper into the building to put additional distance between the rain and me, or to find a room with no windows so that I can't even see the rain, or to keep an umbrella over my head indoors. Practical knowledge is the building: practical knowledge comes with the factivity and the anti-luck condition that, applied to the contents of intention, are all we need for non-defectiveness.²⁴

So the arguments thus far, if successful, show that practical knowledge (of what one is doing, how, and why) is necessary and sufficient for non-defective action by the lights of NEED TO KNOW. But it's worth noting, again, that the kind of defectiveness in question is defectiveness *qua* intentional action. As I emphasized in §2.2.2, it would take significant further argument to assimilate *this* kind of defect to the defects of practical rationality that have been operating in the existing theories about knowledge norms and shiftiness. For all I have said, an action that is non-defective *qua* action, which is to say one that is knowledgeable in the way NEED TO KNOW requires, might still be defective in some other way, or any number of other ways: it might be less than satisfactory aesthetically, or perhaps morally or prudentially or as a matter of etiquette, or, indeed, in terms of practical rationality. (Or, indeed, vice versa.) Again, as I said above, I have not argued that NEED TO KNOW identifies the *only* way in which an action can be defective.

All together, the view we've been building up to entails that intentional actions are subject to a biconditional knowledge norm. In its most general form, it would be something like: an intentional action is non-defective

24. Again, see my (ms) for the details of the argument for NEED TO KNOW.

qua intentional action if and only if it instantiates practical knowledge, the agent's special knowledge of what she is doing, how, and why. But that's not quite the same kind of thing as the norms that have been bandied about in the epistemology literature I've mentioned. Those norms were more like this one, from Hawthorne & Stanley (2008, 578):

HAWTHORNE & STANLEY'S PRINCIPLE: Where one's choice is *r*-dependent, it is appropriate to treat the proposition that *r* as a reason for acting if and only if you know that *r*.

Does NEED TO KNOW entail anything like that? Yes.

The conception of practical knowledge as the special knowledge an agent has of what she is doing, how, and why makes it natural to think of intentions as structured with three 'slots' that must be addressed: One intends to ϕ by means *m* for reason(s) *r*.²⁵ The three slots reflect the kinds of contents needed for an agent to have practical knowledge. Combining the very general formulation of the knowledge norm entailed by NEED TO KNOW and stated above with this view of the content of intentions allows us to reformulate the general statement a bit, to say:

KNOW IT ALL: One can non-defectively ϕ by means *m* because *r* if and only if one has practical knowledge that one is ϕ -ing by means *m* because *r*.

The practical knowledge aimed at in every intention is knowledge of the whole complex. But to have that, one needs to know facts about one's cir-

25. In the full argument for my view in my (ms), I take care to leave open the possibility of acting intentionally for no particular reason, and of basic intentional actions, which is to say actions which require no means (beyond, we might say, the intention itself). But as discussed in more depth there, because of the interdependence of the various (potential) parts of an intention—the way that the question of what counts as ϕ -ing as the agent intends, and questions about why she aims to ϕ , and questions about how she is to ϕ , are all bound up together—it makes sense to understand all intentions as in a certain sense *asking about* reasons and means, even if those slots may sometimes be left unsaturated.

cumstances, one's aims, one's reasons, one's means; knowledge of all these kinds is what enables practical knowledge in our sense. As Anscombe said, having practical knowledge means being able to answer a host of 'why?' questions, in several different ways.²⁶ Thus practical knowledge is constituted by, and requires, knowledge of a bunch of propositions—things like:

- I am ϕ -ing.
- I am ϕ -ing because r .
- r .
- I am ϕ -ing by means m .
- m is a means to ϕ -ing as I intend.
- s is a step in means m , which is a means to ϕ -ing as I intend.

And so on. Therefore, KNOW IT ALL has many corollaries; we derive them by extracting the components of the intention from the complex as contained in KNOW IT ALL, to capture all this necessary component knowledge. So one can non-defectively ϕ because r if and only if one can know that one is ϕ -ing because r , which entails also that one can non-defectively so act if and only if one knows the proposition r that is the reason. Why is that last step true? Because 'because' is factive.

But one might object: Why can't r be my reason even if I don't know it? Indeed, even if I'm wrong about whether r ? This seems like an ordinary occurrence, after all. (I'm not saying that I called my mother for her birthday yesterday even though it wasn't her birthday, but, ahem, one might.) Since 'because' is factive, these won't be cases of acting because r , but it seems as though they could be cases of having r as my reason without knowing it. We can call these cases of acting *on the ground that r* ; the 'on the ground

26. See Anscombe 1957, 9ff.

that' formulation is not factive, as Dancy (2000, 131ff) notes. So why can't I know what I'm doing in the way KNOW IT ALL requires just as long as I know that I'm acting on the ground that r ?

The answer is that when you take r as your reason for acting, you have to take yourself to be acting *because* r , or aiming to, even though you may in fact be mistaken about that. Why? Because acting (merely) on the ground that r leaves open whether r , so to take yourself to be acting (merely) on the ground that r is in some sense to take the truth of r to be irrelevant. But taking something as your reason, *acting for that reason*, can't involve indifference to the truth of the reason. A reason is something that will guide and shape your action as part of your intention, but it can't do that if you take no stand on its truth, and indeed if you are not in an important sense *invested* in its truth.²⁷

Consider: If you ask me why I'm cancelling my hike, and I answer that it's raining, telling me that it's actually not raining after all is a way of telling me that I'm not doing what I intend, i.e., cancelling because it's raining. I won't respond that "Well, all that matters is that my reason for cancelling was that it was raining, never mind that I was wrong about that!" On the contrary, I am likely to lament having cancelled my hike *for no reason*. I was mistaken about what I was doing, to my chagrin. To take something to be a reason to ϕ is to take it to matter whether r , either because the truth of r bears on the desirability of ϕ -ing, or because the truth of r counts in favor of ϕ -ing in some particular way rather than another. If r is your reason, then it matters for guidance. So while I can be mistaken or uncertain whether r , I cannot be agnostic whether r . That is why I must take myself to be acting because r , which entails r , and not merely on the ground that r , which does not entail r .²⁸ And if I take myself to be acting because r , then I will be

27. My thinking about this question draws on Hyman (1999), as well as Gibbons (2001).

28. Of course, if I am uncertain whether r , that very uncertainty is likely to shape my action; sometimes I will think acting on r under uncertainty is not worth the risk,

mistaken about what I'm doing, in the way prohibited by KNOW IT ALL, if I fail to know r .

To come at the question another way: what we are talking about in talking about my acting on the ground that r (without taking myself to be acting because r) is intending to ϕ because *I (merely) believe that r* . But in a case like that, it is not r that is my reason, but my mental state. And if I incorporate my belief that r into my behavior in this way, *while agnostic about the truth of r* , so that there is not any sense in which what I aim at is to act *because r* , what we'll have is not a case of acting for the reason that r , but of acting for the reason that *I merely believe that r* —in which case it will not be r that guides my action, but my belief (and a strangely self-consciously-mere belief, at that).²⁹ But in any case these will not be cases in which I take r as my reason for acting, so they are not counterexamples to the claim that what practical knowledge requires includes knowledge of the propositions that are your reasons—even though it is true that we are in fact often mistaken about purported facts we take to be reasons, and we often explain our actions (especially retrospectively, after things have gone awry) in terms of things we took to be reasons that were not actually the case. The point is that *in acting* we cannot take our reasons to be other than facts, and so we must know those facts if we are to know what we are doing.

or sometimes I will alter my action so that I am acting on a weaker proposition, like probably r or possibly r —as when I bring handouts for the conference even though I'm planning to use slides, because I think it's possible I won't have the right dongle to connect my laptop to the projector. This makes perfect sense on my view.

29. It's not obvious that this odd state would count as believing r . What would it mean to believe that r while agnostic about the truth of r ? But if I can believe r in this way, if cases like this exist, I think they will be rather like certain cases in which one acts from emotion without taking the emotion to be a reason for the action, along lines discussed in Hursthouse (1991). So I might smile out of frustration, even though I don't think being frustrated is a reason to smile; perhaps there are roughly analogous cases for acting from a self-consciously-mere belief.

All of which is to say that one corollary of KNOW IT ALL is that one can non-defectively ϕ for the reason that r if and only if one knows that r , because an agent, in acting, must take herself to be ϕ -ing because r . Any proposition that fills the r slot in an intention will make that intention (as well as the action that is the intention's execution) defective if the agent fails to know the proposition. Therefore:

KNOWN REASONS: One can non-defectively act for the reason that r if and only if one knows that r .

But KNOWN REASONS looks quite similar to Hawthorne & Stanley's principle (henceforth HSP). So let's take stock.

Both KNOWN REASONS and HSP put appropriately/non-defectively acting with r as one's reason on one side of the biconditional, where the other side is about knowing r . But I have suggested that there is at least some reason to think the kind of appropriateness in question is different.

Before I get into the different senses of appropriateness, though, it's worth noting here a couple of other ways my view is different, and perhaps more appealing, than those in the literature: on my view, the two directions of the biconditional really are inseparable, symmetrical, and on equal footing. Contrast Hawthorne & Stanley (and Fantl & McGrath), whose sufficiency direction seems both less robustly supported when compared with the necessity claim, and not symmetrical with it, since only the sufficiency direction is qualified. My cleanly symmetrical biconditional is so precisely because both sides are generated by one and the same argument. But Hawthorne & Stanley (and Fantl & McGrath) have to provide a separate rationale for the (restricted) sufficiency side.

Why think that the sufficiency direction is less well-supported than the necessity side? Recall that the necessity direction rests on the Criticism cases: it seems clear that acting on the basis of p without knowing p makes one subject to reproach. But sufficiency is not supported in this way. For Hawthorne & Stanley and Fantl & McGrath, I said, it seems that the sufficiency direction rests on the 'it's weird to combine a claim that you

know p with a declaration that you ought to do something that is not what you ought to do conditional on p ' move. But that move seems not really to hinge on the *knowledge* part of the picture; it's the definition of acting as if p/p dependence in terms of a (rational) preference ordering based on p that does the work. While the faultiness of failing to prefer as if p when one is in a p world does mean that if you know p you ought to act as if p in the relevant sense, on pain of irrationality, it likewise means that you at least sometimes ought to act as if p if you merely believe p , or only if you are certain that p , also on pain of irrationality. All of which is to say that Hawthorne & Stanley (and Fantl & McGrath) have to do additional work, and via a strategy that seems not really connected to the project of constructing a *knowledge* norm, to get something that falls out automatically on my view.

But those are subtleties; let's talk about the flavor-of-appropriateness question. In HSP, I have noted, it is natural to understand the kind of appropriateness in question as having to do with practical rationality. Hawthorne & Stanley (2008, 572) talk about how it would be wrong to take *this ticket will lose* as a premise in practical reasoning about whether to sell a \$1 ticket in a 10,000-ticket lottery with a \$5,000 prize for one cent. In the case, the talk of what is appropriate to take as a premise echoes Weatherson's (2012) concern with what to write on a decision table; the specifics of the odds and the dollar values suggest that expected utilities of a standard sort are what determines if the sale is a bad deal. In KNOWN REASONS, the kind of appropriateness (called, in my preferred formulation, non-defectiveness) in question is about achieving practical knowledge. One must know one's reasons because reasons figure in intentions, and practical knowledge is knowledge in intention. To see whether one errs by the lights of KNOWN REASONS, one does not ask what the expected utilities are; one asks if p is a constituent of the intention.

So there are differences between my principle and previous proposals, despite surface similarity. And perhaps the most important difference, of course, is that my principle is supported by a ground-up argument, rather than just intuitions about cases. So if I can explain everything they can, and my

proposal has advantages both subtle and stark, it looks like my view is a better option.

So let's consider some cases and compare results.

2.4 COMPARING RESULTS IN CASES

My principle will give us the 'expected' results in many cases; when my view is at odds with its rivals, I'll suggest, my view gives a better picture. Let's start with the criticism cases.

2.4.1 HSP VERSUS KNOWN REASONS IN CRITICISM CASES

In fact, HSP and KNOWN REASONS will often give the same intuitively plausible results. For example, in criticism cases like *Restaurant*, they will agree that Hannah errs if she takes *the restaurant is down the street to the left* as her reason without knowing it to be so. My norm and the proposals from the interest-relativists will agree that agents who act on the basis of *r*, or take *r* as their reason for acting, go wrong if they do so without knowing *r*.

But notice that the two norms will give different results if we vary the case a little. Suppose we revise *Restaurant* so that Hannah's reason isn't *the restaurant is down the street to the left*, but rather *the restaurant is probably down the street to the left*. Hannah could well know that, even if she doesn't know that *the restaurant is down the street to the left*. But if we make this change—call the new case *Restaurant**—then the enforcer of HSP has no complaint to make. Hannah has taken something she knows as her reason.

What does KNOWN REASONS tell us? It will agree with HSP that Hannah acts appropriately in *Restaurant**. However, there is more to say: KNOWN REASONS is a mere corollary; it is derived from KNOW IT ALL. And if we fill in Hannah's intention, so that what she intends is *to get to the restaurant on time, because the reservation is time-sensitive*, then if her intended means does not in fact secure their on time arrival, her intention is defective. And

if she knows only that the restaurant is probably down the street to the left, it is lucky, in the way that rules out Hannah's being in the good case, if they get to the restaurant on time. Her intended means is *gappy*; it leaves it open whether or not that means is sufficient to secure the intended outcome. She needs a favor from the world.

Of course, it is often the case that we have to act on fall-back propositions like *probably p*. It is frequently the best we can do. But what my view brings out, via the combination of KNOWN REASONS and KNOW IT ALL, is both what Hannah does right and the ways in which she would wish for better. Even when we don't or can't, limited and fallible creatures that we are, know the full-strength *p*, and we therefore, because of other features of our situation, decide to act on *probably p* because it is the best we can do, we remain sensitive to the way in which the full-strength *p* would be better. Thompson (2011, 209–10) makes a point in this spirit about the carbon copier: he says that if one really were in that case, one would check along the way to see how it was going, not write the whole page and then find out. My view explains this; we often take whatever degree of practical knowledge we can get, while still recognizing that more would be better. And if we execute an overall defective intention when a non- or less-defective one is available, we err. HSP does not give us the materials to see all this.

Now, a partisan of HSP might reply in one of three ways. First, she might say that in fact she can capture the residual defectiveness in a case like *Restaurant**, via the requirement of *p*-dependence: maybe the ranking of worlds preferable if *p*, when *p* is the proposition that the restaurant is probably down the street to the left, will be such that turning left is not preferable. But I think this is not promising, and in any case, as I said above in talking about the asymmetry of HSP in §2.3, going this way would mean that all the interesting work is being done by the requirement of *p*-dependence, not the requirement of knowledge, and indeed would seem to apply to a wide range of attitudes, including mere belief and robust certainty.

Alternatively, she might say that the residual defectiveness is not part of

the phenomenon in which she is interested; it's a separate question, like questions about æsthetic value or etiquette. Again, I do not think this is promising. This residual defectiveness has to do with the way an executed intention does or does not serve to secure the agent's aimed-at outcome, as is clear on my view. This is not, intuitively, an extraneous consideration to the evaluation of action. Indeed, in the full argument for my view in my (ms), the connection between practical knowledge and skill makes it clear that this way of thinking is central to our evaluation of action: skillful actions are ones in which the agent is better able to shape the world to her will. Sometimes she will desire that the action be elegant, or polite, and if she is skilled she will be able to make it so. But just noticing the structure of that thought—the way that skill, the ability to manifest control—is a means to the agent's ends, which might include things like æsthetic or other kinds of value, shows that evaluation in terms of practical knowledge is not on a par with those other forms of evaluation.

Finally, the HSP partisan might simply deny that there is any residual defect here. But my description of *Restaurant** does seem to capture something significant, and of course my view underpins the case with theory built from the ground up. So why should we take the HSP partisan's side?

To summarize, while HSP and KNOWN REASONS give the same results in a range of cases—calling foul anytime an agent takes *p* as her reason without knowing *p*—they can disagree. And when they do, I am suggesting, my view, combining KNOWN REASONS and KNOW IT ALL, gives a more illuminating account. In those cases, either the HSP partisan can't capture what's going on in *Restaurant**, or she can only do so by making use of machinery that is external to the project of explaining why knowledge is important for action—whereas my view makes the connection organic. My view does better.

2.4.2 HSP VERSUS KNOWN REASONS IN ENCROACHMENT CASES

What about encroachment cases? I have not yet said anything about those. But the shiftiness claims that pairs like *Bank (low stakes)* and *Bank (high stakes)* are supposed to motivate are the whole point for Fantl & McGrath,

and Stanley (though perhaps not Hawthorne), and Weatherson.

The first thing to say is that KNOW IT ALL and KNOWN REASONS don't obviously say anything at all about these cases. Remember, what we test for with my norms is whether the agent knows what she is doing, and how, and why. But what drives the verdicts about encroachment in the literature is a claim about practical rationality; the claim is that, roughly, the expected utilities won't come out right when the stakes are high. Since my view is not based on such verdicts, it seems reasonable to say it makes no predictions about encroachment.

This is already an interesting result: apparently, we can have a strong, biconditional knowledge norm for action that does not entail shiftiness. Because stakes seem irrelevant to whether or not one knows what one is doing, there's just no basis for a claim of encroachment. The indirect route to encroachment—on which we move from a verdict that the agent should not ϕ , plus a norm like HSP, to the claim that the agent does not know some p —is not supported by my flavor of norm.³⁰

This creates a new argumentative burden for those who have wanted to argue for the entailment: they will need to demonstrate conclusively that they are not simply mistaking my non-entailing norms for something that does have such a consequence, especially since my norms capture at least as many of the intuitive results (and are supported by a ground-up argument). Of course, it would still be open to those who like encroachment to retreat to the direct reading of these cases; if the knowledge norm plays no part in the claim of encroachment, it doesn't matter if there are two possible flavors of knowledge norm. Nothing I have said obviously militates against that view.

But such a retreat is costly. First, because the direct reading, in foregoing

30. For reference, the discussion of indirect vs. direct encroachment is found in §2.1.2 above.

the explanatory component that the indirect reading offers, is less appealing just in virtue of lacking that explanatory component. Second, since the direct reading relies essentially exclusively on intuitions about loss of knowledge in high-stakes cases, it is both dialectically weaker and explicitly vulnerable to evidence that those intuitions are not as robust or as clear as the direct reading supposes—and recent empirical work, such as Schaffer & Knobe (2012) and Rose, Machery, et al. (2017), provides reason to worry about relying on these intuitions.

Of course, we might wonder: Is there anything in my view that actively rebuts claims of encroachment, rather than just failing to support them?

Let's start thinking about that by exploring what my view might offer if one wanted to say that Hannah actually can know the relevant p in both the low- and high-stakes cases. (After all, as I said, one would naturally think that stakes are irrelevant to whether Hannah knows what she is doing.) Consider a slightly fleshed-out version of the pair of bank scenarios.

This time, suppose that Hannah is in the high-stakes case and intends, plausibly enough, *to make the deposit on Saturday because the bank will be open then and the line will be shorter and the check will still clear before Monday*. Now suppose as a companion case that Hannah is in the low-stakes case, and, also plausibly, intends *to skip the line tonight because either the bank will be open tomorrow or, if something bizarre happens, the deposit can go in on Monday*. Now, it's true that Hannah might, as a first pass, describe both these intentions with the same words. If you asked her what she was going to do about making the deposit, she might, in both cases, answer, "I'm going to wait till Saturday to avoid the lines." But what I want to suggest is that we shouldn't take that to be the end of the story.³¹

31. cf Fara (2013), for an extensive discussion of this kind of underspecification in the case of a similar attitude, desire. My point here can be read largely as a translation of her basic idea to the intention case. There is a good deal more to say about just what intentions are like on my view, but I can't give the whole picture here; it is the project of a further paper.

The differences between the two intentions are subtle, but what I hope to suggest is that as a matter of psychological fact we may, generally speaking, configure our intentions somewhat differently depending on the stakes, and in ways that matter. So high-stakes Hannah is in fact taking the bank's being open on Saturday as her reason, whereas low-stakes Hannah is taking something weaker, like *the bank is probably open on Saturday* as her reason. If that's right, then high-stakes Hannah, but not low-stakes Hannah, errs if she fails to know that the bank will be open on Saturday. Remember: what my norms test for is whether the intention is such that the agent can have *knowledge in intention*, so what is required is knowledge of what one is doing, how, and why.

To come at the thought another way: It is quite natural for some kind of cost-benefit analysis to be among our reasons for acting. So even if we haven't actively or explicitly considered the question, if asked one of Anscombe's why questions about the costs and benefits of what we're up to, we will answer. Ask low-stakes Hannah why she is risking finding out that the bank has changed its hours so that she'll have to wait till next week to make the deposit, and she'll say that it's worth it to avoid the lines. But one thing that we might expect of Hannah is that if she's in the high stakes case, she will in fact, as part of her intention, aim to be ruling out even very unlikely scenarios such as the bank's changing its hours—calling the case high-stakes presupposes that the cost is substantial and to be avoided. If her intention is configured that way, then it will be a *different intention* than low-stakes Hannah's intention, which does not aim to rule out this unlikely possibility, and there will therefore be different ways for it to be defective, because different propositions will need to be known.

One way to see this is to think about the range of questions we might ask Hannah. I said that low-stakes Hannah, we can suppose, will not be configuring her intention so as to rule out a change of hours. But if you ask high-stakes Hannah "Why are you risking the bank's being closed because it has changed its hours?" I suggested we can expect her to say "I'm not—I know they're open on Saturday." But if you ask even high-stakes Hannah

something a bit different, say, “Why are you risking the bank’s being closed because of an alien invasion?” you will likely find there are wild possibilities like this that her intention is not built to rule out—“Well, if we have an alien invasion I’m thinking I’ll have bigger problems than whether the rent clears.” In other words, on my view high-stakes Hannah will be relying on a stronger proposition than low-stakes Hannah, but it’s very likely that neither is relying on a strict, literal reading of ‘the bank will be open on Saturday.’ This is what I mean when I say that the way we describe our intentions will be underspecified. High-stakes Hannah and low-stakes Hannah will probably have intentions that differ in quite a few respects, including what claims they make about whether the bank will be open. (And whether the subway will be running on Saturday, and whether they will get caught up binge-watching *The Handmaid’s Tale*, and a host of other potential obstacles to getting the check deposited on Saturday.) This is true even if they might give the same (underspecified, shorthand) answer to what their plans are.

But this would just be a contingent psychological fact about Hannah. If she were, instead, a financial thrill seeker, she might intend to take the risk even when the stakes are high, in which case her action would not (or at any rate, might not) be defective *qua* action.³² This is thus quite a different picture than one in which the epistemic standard for knowledge varies with stakes.

32. I say ‘might not’ because we haven’t yet said quite enough about Hannah’s total cost-benefit picture. Keeping in mind how open the available ‘why?’ questions are, it might be that even though Hannah intends to risk the bank’s being closed (because she’s a thrill seeker), she might also not intend to, say, risk losing her job—but it might be that if the bank is closed on Saturday so that her rent check bounces on Monday, the mark on her credit caused by the bounced check will show up on a random credit report performed by her employer. Hannah might therefore have an intention that is not fully consistent—because she is acting on the ground that *p* and *q*, where *p* is that *I’m taking the thrilling risk of having the bank be closed* and *q* is that *I’m not risking losing my job*, but *p* and *q* cannot, as it happens, both be true—and this inconsistency means that her intention can never be knowledge, which means it is defective by the lights of NEED TO KNOW.

There's a lot of machinery here, but what I'm trying to get at is really two points. The first is that agents very likely tend, as a matter of psychological fact, to construct their intentions differently depending on things like stakes, even if a casual statement of an intention would not reflect this.³³ The second is that what will count as defective intentional action by the lights of my view is a matter of what propositions actually do figure in an intention, and of certain requirements of consistency among these constituent propositions. Because the aim is knowledge in intention, the various propositions about means, reasons, etc. must be known, and they must fit together in such a way that the whole complex can itself be known.

Combined, I want to suggest, these two thoughts mean that we can make good sense of many encroachment cases—sometimes getting the same results as the interest-relativists, and sometimes getting different, but (I have argued) better, results—*but without needing to claim that knowledge is shifty*. In high-stakes cases, agents will tend to construct their intentions so as to rule out the possible bad outcomes, which will require knowledge of stronger propositions (p rather than *probably* p , for example) or more propositions (because one might intend finer control, and so need more detailed knowledge of means, for example). It will thus indeed normally be harder to know what one needs to know when the stakes go up; but this will be only a matter of contingent psychological fact, and the explanation is not that what it takes to know a given proposition varies with one's practical situation, it's that what propositions must be known varies with one's practical situation.³⁴

33. Again, see Fara 2013.

34. My view is thus compatible with something like the insensitive invariantism proposed in Williamson (2005). But where Williamson says that it may be that when stakes go up we simply, as a psychological matter, want to know that we know p , or know that we know that we know p ; I am making a more general claim that what we want to know, usually, when the stakes go up, is just *different stuff*, which will often include the iterations of knowledge he mentions, but also will likely include stronger and/or more propositions of various kinds, as discussed here.

All of which is to say that on my view, a plausible empirical regularity in human intention formation in response to changes in stakes is likely, in combination with the underspecification of intentions and their contents, to generate patterns that look very much like those predicted by interest-relativists and other shifty theorists. Depending on how we fill in the cases, we will often get the same intuitions about propriety. We can vindicate the sense that it's often harder to know what you need to know to act properly when the stakes are high—but in a way that has very different theoretical upshots.

It's worth considering one more example, to bring out the key thought that we can expect agents in high-stakes situations to, generally, tailor their intentions so as to avoid the possible very bad outcomes.³⁵ Suppose I am a neurosurgeon operating on someone's brain, so that if I make a mistake I am likely to cause severe brain damage. It will be natural for me to choose my intended means so as to *rule out* such mistakes; I'll add various double-checking steps, I'll avoid coffee if it might make my hands shake, etc.

But if we suppose instead that I am a neurosurgeon trying out an idea for a new surgical technique in a computer simulator, it will then be unsurprising if I intend to omit some of those steps. This difference in my intention changes what facts I need to know: when I'm performing the real surgery, I need (if I am a normal, responsible neurosurgeon) to know propositions such as *I confirmed that the patient's tumor is in the right cerebral hemisphere* and *my hands are not shaking*; if I'm just trying an idea out in the simulator, maybe part of what I want to find out is if having coffee beforehand will make my hands shake too much to pull it off, so that I would not need to know that *my hands are not shaking*.

What needs to be known, on my view, is a matter of what is required for the intention to be a candidate for knowledge: so anything that figures in

35. Even though this is, as I have said, a contingent feature of the agents' psychology, which we can see by reflecting on our alternative Hannah, the financial thrill seeker.

the intention—as a reason, or as a means, or as an end—must be known, and the various components must not be inconsistent or otherwise such as to be jointly incompatible with knowledge of the whole. So when agents tend to tailor their intentions in the face of high stakes so as to rule out the possible very bad outcomes, that will change what propositions must be known—and, usually, by requiring either knowledge of stronger propositions, or knowledge of more propositions, or both, thereby increasing the agent’s epistemic burden.

But even though the picture I’m proposing gives us reason to expect that in most cases, agents will react to increases in stakes by altering their intentions in ways that make it more difficult to know what they must in order to have practical knowledge, the view does *not* support the claim that this increased epistemic burden *necessarily* comes with increases in stakes (because we can construct cases where it does not, along the lines of the financial-thrill-seeking Hannah, above), or that knowledge itself is interest-relative (because the reason it is often harder to know what you need to know when stakes are high need not have to do with a shifting standard for knowledge, but can rather be about changes in what contents must be known).

2.4.3 A CAVEAT: NOT JUST CASES

I have said that the arguments for shifty epistemology proceed largely on the basis of intuitions about these two types of cases, and that this leaves the shifty theorists at a disadvantage when faced with my view, which is supported by a ground-up argument. But I should note that Fantl & McGrath (2009, but especially 2012) do attempt to provide a principled argument for shiftiness. Unfortunately, it does not help them with the arguments of this paper.

Their argument (Fantl & McGrath 2012, 64–72) is really more of a schema, as they take it that there are any number of ways of filling out the premises, but the key idea, they say, is that if knowledge is compatible with something less than certainty, *and* knowledge requires what they call “actionability,” then there will be pairs of cases like the high- and low-stakes Bank cases,

where changes in stakes affect actionability, and so whether one knows, in basically the way I describe in discussing the indirect reading of the Encroachment cases in §2.1.2: the expected utilities won't work if knowledge doesn't require probability 1.

At the heart of the argument is the claim that *no matter how you understand actionability*—no matter what the specifics are of the principle you use to link knowing p to being able to (non-defectively) act on p —you will find that whether p is actionable in a particular case will depend on stakes.

But that is precisely what I deny: my view gives us knowledge norms that serve perfectly well as statements of actionability, but they do not combine with fallibilism to entail shiftiness. What does the work for me is not degrees of certainty and expected utilities, but the fact that knowledge is factive and rules out relevant kinds of luck. So Fantl & McGrath's principled argument fails, because there is a way to understand actionability that does not give them what they need. (It's worth noting that their argument actually still requires the case judgments, too.)

2.5 CONCLUSION

All told, this paper does several things. First, and crucially, it connects—for the first time—discussions about the link between knowledge and action that have taken place in Anscombean action theory and in recent epistemology, and shows that action theory can help illuminate the epistemological debate. To that end, it explicates some knowledge norms that follow from my Anscombe-inspired view of the nature of intentional action, and shows how those norms can deliver satisfying results in criticism and encroachment cases, and often more illuminatingly than previously proposed norms—so my view can honor the intuition that, for example, an agent who takes r as her reason for acting, but does not know r , has gone wrong, concurring with the shifty theorists but also offering an explanation; and my view allows us to say more about why certain cases where an agent does know her reason will still be less than ideal.

Finally, it serves at least to bring into question the motivation for claims that knowledge is shifty. Because the main basis for earlier discussions has been intuitions about cases, and my view captures the cases at least as well, while also being supported by a ground-up independent argument, the burden is now on those who want to argue either for knowledge norms based in practical rationality, or for the interest-relativity of knowledge, to explain why those views should still be tempting in light of what I've said here. I have not offered a knock-down argument against practical-rationality-based norms, or against shiftiness. But I have eroded the rationale for those views, and in ways substantial enough, I take it, to be debunking.

It may have seemed, till now, that recognizing the strong connection between knowledge and proper action required going in for shiftiness. I hope to have offered good news to those who find that a high price to pay. You can intentionally acknowledge the importance of knowledge to acting well without having to get all shifty about it. You just have to know what you're doing.

REFERENCES

- Anscombe, G.E.M. (1963) *Intention*. Cambridge, MA: Harvard University Press.
- Cohen, S. (1999) "Contextualism, skepticism, and the structure of reasons," *Philosophical Perspectives* 13: 57–89.
- Dancy, J. (2000) *Practical Rationality*. New York: Oxford University Press.
- Davidson, D. (2001) *Essays on Actions and Events*. New York: Oxford University Press.
- Fantl, J. & McGrath, M. (2002) "Evidence, pragmatics, and justification." *The Philosophical Review* 111(1): 67–94.
- . (2009) *Knowledge in an Uncertain World*. Oxford: Oxford University Press.
- . (2012) "Arguing for shifty epistemology." In *Knowledge Ascriptions*, ed. Brown, J. & Gerken, M. Oxford: Oxford University Press.
- . (2013) "Practical matters affect whether you know." In *Contemporary Debates in Epistemology*, ed. Steup, M., Turri, J., & Sosa, E. Oxford: Wiley Blackwell.
- Fara, D.G. (2013) "Specifying desires." *Noûs* 47(2): 250–272.
- Gibbons, J. (2001) "Knowledge in action." *Philosophy and Phenomenological Research* 62(3): 579–600.
- Hampshire, S. (1967) *Thought and Action*. New York: Viking.
- Hawthorne, J. & Stanley, J. (2008) "Knowledge and action." *The Journal of Philosophy* 105(10): 571–590.
- Hursthouse, R. (1991) "Arational actions." *The Journal of Philosophy* 88(2): 223–246.
- Hyman, J. (1999) "How knowledge works." *The Philosophical Quarterly* 50(197): 433–451.
- Jaques, A.E. (forthcoming) "Knowing and doing, well and good." *Oxford Studies in Metaethics*.
- . (ms) "Where there's a will, there's a way of knowing."

- Korsgaard, C.M. (2009) *Self-Constitution: Agency, Identity, and Integrity*. Oxford: Oxford University Press.
- Railton, P. (1994) "Truth, reason, and the regulation of belief." *Philosophical Issues* 5: 71–93.
- Rose, D.; Machery, E.; et al. (2017) "Nothing at stake in knowledge." *Noûs* 00(0): 1–24.
- Schaffer, J. & Knobe, J. (2012) "Contrastive knowledge surveyed." *Noûs* 46(4): 675–708.
- Setiya, K. (2008) "Practical knowledge." *Ethics* 118(3): 388–409.
- . (2012) "Knowing how." *Proceedings of the Aristotelian Society* 112(3): 285–307.
- . (2013) "Murdoch on the sovereignty of good." *Philosophers' Imprint* 13(9): 1–21.
- . (2014) "Intention, plans, and ethical rationalism." In *Rational and Social Agency: The Philosophy of Michael Bratman*, ed. Vargas, M. & Yaffe, G. Oxford: Oxford University Press.
- . (2016) *Practical Knowledge: Selected Essays*. Oxford: Oxford University Press.
- Stanley, J. (2005) *Knowledge and Practical Interests*. Oxford: Oxford University Press.
- Thompson, M. (2008) *Life and Action: Elementary Structures of Practice and Practical Thought*. Cambridge, MA: Harvard University Press.
- . (2011) "Anscombe's *Intention* and practical knowledge." In *Essays on Anscombe's Intention*, ed. Ford, A., Hornsby, J., & Stoutland, F. Cambridge, MA: Harvard University Press.
- Weatherston, B. (2012) "Knowledge, bets, and interests." In *Knowledge Ascriptions*, ed. Brown, J. & Gerken, M. Oxford: Oxford University Press.
- Velleman, J. D. (1989) *Practical Reflection*. Princeton: Princeton University Press.
- . (2000) *The Possibility of Practical Reason*. Oxford: Oxford University Press.

———. (2009) *How We Get Along*. Cambridge: Cambridge University Press.

Williams, B. (1973) *Problems of the Self*. Cambridge: Cambridge University Press.

Williamson, T. (2005) “Contextualism, subject-sensitive invariantism and knowledge of knowledge.” *The Philosophical Quarterly* 55(219): 213–235.

3 SHACKLING

Here's a thought: When you're a marginalized person, you tend to be judged less knowledgeable than you are, and the conceptual tools people use to produce and structure knowledge tend to mask important facts about you and your experience. That's *epistemic injustice*, and it's everywhere. (Or so Fricker (2007) argues.)

Here's another: Acting well requires knowledge; if you don't know what you're doing, and how you're doing it, and why, something is going wrong. (Or so I argue elsewhere.)¹

I think both those thoughts are true, and furthermore that they interact in an important way—or so I'll argue in what follows. Together, they entail that to be denied the status of knower is to be denied the possibility of acting well: so epistemic injustice does violence to the will.

If one cannot act well, either one must not act, or one must act defectively. Those denied knowledge are thus in a double bind, forced to choose between objecthood and wrongdoing. I call this *shackling*, because the word captures the two sides of the predicament: shackles both constrain movement and mark one out as a criminal. That's the phenomenon I'll explore in this paper. As I'll show, given what action is like, epistemic injustice leads to agential injustice: harms to agents *qua agents*.

1. See my (ms a).

I'll begin with some background: in §1, I'll lay out my preferred way of thinking about epistemic injustice, and in §2 I'll sketch my view about the nature of intentional action and its connection to knowledge. The heart of the paper begins in §3, where I'll connect these two lines of thought, setting out in abstract terms the kinds of interactions we should expect to see—the structure and mechanics of shackling. In §4 I'll discuss some cases that demonstrate the ways shackling works in our world, before offering some concluding thoughts about what it all adds up to in §5.

3.1 WHAT IS EPISTEMIC INJUSTICE?

The phenomenon of epistemic injustice was famously described by Fricker (2007). She identifies two kinds, each of which is, she says, a distinctively epistemic wrong, something 'done to someone specifically in their capacity as a knower' (Fricker 2007, 1):

Testimonial injustice (according to Fricker): an unjust reduction in the credibility of a speaker due to prejudice in the hearer

Hermeneutical injustice (according to Fricker): an unjust reduction in the intelligibility (to herself and/or someone else) of facts relating to a person who is a member of a marginalized group, due to gaps in the community's interpretive resources

As stated, these definitions raise a lot of questions, even beyond "Can you give me an example?" One might well wonder: What kind of injustice is this supposed to be? Where and when does the injustice occur? Is it really only a matter of individual exchanges? How can a hearer's failure to believe be a harm to a speaker, and *wrongful*? What's especially wrongful about failure to believe that's caused by prejudice? Is hermeneutical injustice only a matter of a total absence of interpretive resources? When and why is it *unjust* for someone not to have interpretive resources that would be helpful, rather than merely unlucky? (Are we to suppose that every worker who lived before Marx's time was *unjustly* deprived of his insights? How could

that be?) Are the harms here purely instrumental? If so, how are they distinctive? And so on.

In what follows, I will be working with a conception of epistemic injustice that diverges from Fricker's in a few important ways. While I'll follow her in discussing two broad types of epistemic injustice, I'll understand each of them rather differently. Going my way will not only give us a better grip on some important phenomena (or so I'll suggest), it will also help with the questions above—either answering them, or making them less pressing. So rather than taking up the questions now, I'll set out the way of thinking I propose, highlighting the differences with Fricker and giving examples to make clear what the target phenomena are supposed to be, so that we can return to the questions with that view in hand.

I should say that I'm not going to offer a detailed argument that my view is *the correct* view of epistemic injustice, *as opposed* to Fricker. First, because I think it's plausible to understand what I offer as a friendly amendment to and extension of her view. Second, I really just mean to describe something whose existence in the world should be obvious (by the time I'm done) and then *stipulate* that I'm going to call that thing 'epistemic injustice.' As we'll see, I think my view can deal with or avoid some of the thorny questions that arise for Fricker, and I take it that's a reason to think I'm onto something. But I'm not interested in fighting about what should be the One True Statement of epistemic injustice; for all I'll say here, there may be many useful views in this neighborhood that are suited to different purposes.

3.1.1 THE EPISTEMIC ECONOMY

The first key difference between my conception of epistemic injustice and Fricker's account is that where Fricker takes epistemic injustice to be a matter of wrongs done to and by individuals, occurring in particular interpersonal interactions, I think it is, in the first instance, a *structural* phe-

nomenon.² Fricker locates testimonial injustice in, for example, an event in which a police officer fails to believe a young Black driver who says he is the owner of the car he is driving through a white neighborhood.³ She says hermeneutical injustice occurs, paradigmatically, when Carmita Wood is unable to express her legitimate reason for quitting her job to the person in charge of processing her departure, because the concept of sexual harassment has not yet been invented.⁴ While I agree with Fricker that we should expect to see interactions between individuals that flow from and/or instantiate and/or manifest epistemic injustice, I propose that we have a lot to gain by focusing our attention on the structural, rather than the individual, level.

What do I mean by that? I mean that epistemic injustice, as I understand it, has to do with the shape of, and patterns in, the social world.⁵ In particular, I understand epistemic injustice to occur both in and through what I call an *epistemic economy*.⁶

-
2. In this, I take myself to be in harmony with critiques of Fricker's individualistic methodology raised by Anderson (2012) and Haslanger (ms).
 3. See Fricker 2013 (1319) for this case, a more detailed version of a general kind mentioned several times in her 2007, beginning on p.1.
 4. This case, now treated as paradigmatic throughout the literature on epistemic injustice, is discussed at length in Fricker 2007 (149–151). But note that she is explicit that even though the lack of the necessary concept suggests a systematic, structural problem, the injustice only occurs in particular transactions: “the moment of hermeneutical injustice comes only when the background condition is realized in a more or less doomed attempt on the part of the subject to render an experience intelligible, either to herself or to an interlocutor” (Fricker 2007, 159).
 5. When I talk about patterns here, I don't mean to assume that justice is fundamentally distributive, or to rule out a view of justice like Anderson's (1999) democratic egalitarianism. I mean only that the fundamental locus of the problem is at the social/structural, rather than the individual/transactional, level.
 6. 'Economy' talk is not alien to Fricker; she mentions a 'credibility economy' and an 'economy of hermeneutical resources' (Fricker 2007; 1, 30). But she does not pursue

An epistemic economy is composed of various epistemic goods, along with a system of practices involved in the production, reproduction, and distribution of those goods. Epistemic goods, or epistemic resources, can include anything of positive or negative epistemic value; so the epistemic economy will traffic not only in knowledge, but also in what we might think of as the raw materials for knowledge—such as, importantly for my purposes here, conceptual and/or interpretive resources—as well as things like *credibility* that are needed to participate in and benefit from the operation of the epistemic economy. It's natural to think of knowledge as the fundamental good epistemic economies are concerned with, but conceptual tools and credibility are also produced, reproduced, and distributed within such systems, primarily *as a means to* the production, reproduction, and distribution of knowledge. Many kinds of epistemic goods flow through this economy.

The provision of an exhaustive list of epistemic resources is beyond the scope of this paper. But we can get a sense of the kinds of things that list will include. *Evidence* is surely on the list, and I think that certain material objects and even institutions (or access thereto) will also appear—consider *libraries*. But both evidence and libraries suggest that list items will be such *under certain guises*: a library is an epistemic resource insofar as it is a repository of knowledge; it is not an epistemic resource when it's considered as a place to get out of the rain. Given the many different accounts of what evidence is—among which I will remain agnostic—we may need to make similar moves. If evidence is just knowledge, as Williamson (2000, 185) thinks, it's obviously an epistemic resource; if evidence can be fingerprints on a bloody knife at a murder scene, we'll need to say that insofar as those fingerprints help the detective learn the identity of the killer, they are an epistemic resource; insofar as they contribute to a mess that will need to be cleaned up, they are not.

the idea, and never suggests that there is an overarching *epistemic* economy in the sense I'm interested in.

What this means is that some things will, in certain circumstances, be properly called epistemic resources, even though it seems odd to call them such a thing more generally: say, money, or health care. I think this is no great mystery, though. Money and health care are resources of extreme generality, useful and/or necessary for a huge range of human undertakings. They are produced, reproduced, and/or distributed in many human systems. So calling them *epistemic* resources will often be misleading, even if correct, in that it will suggest that the role they play in the epistemic economy is their *primary* role. It may well be that to build the exhaustive list of epistemic resources, we'll need a definition as broad as 'knowledge, plus anything instrumental for knowledge'—which we'll then have to supplement with lots of explanations of the various guises under which particular items do or do not count.

But as I said, that project would take us off track. In what follows, I'll talk mostly about two kinds of epistemic resources other than knowledge: conceptual/interpretive resources, and credibility. What makes those things epistemic resources? Let's start with credibility.

As mentioned above, credibility is needed for full participation in the epistemic economy; as Daukas (2006, 109) emphasizes, "epistemic cooperation requires trust," and we base decisions about who to trust on assessments of others' epistemic competencies. 'Credibility,' as I'll use the term here, is close to Daukas' 'trustworthiness.'⁷ As she sees it, this trust is epistemic because it has to do with assessments of how closely someone approximates an ideal of behaving "(when contextually appropriate) as though her epistemic status is S if and only if her epistemic status is S" (Daukas 2006, 111). However, as Daukas also notes, the trust required for full participation has a moral component as well; we assess sincerity, benevolence, etc.,

7. Daukas' discussion brings out the way that these assessments can contribute to and constitute marginalization; I'm indebted to her discussion, which predates Fricker's book.

along with epistemic habits in deciding who to trust. So if we're roughly following Daukas, we can see that credibility, or trustworthiness, has an epistemic side—assessment of whether someone behaves in accordance with her actual epistemic position—that is admixed with moralized assessment. (This hybrid quality will figure in some of my discussion later.) Credibility is an epistemic resource because it is a prerequisite for participation in the epistemic economy. We do not engage with others in the giving and receiving of knowledge when we think they are untrustworthy; we tend not to include them in the systems of production and reproduction of knowledge, and we tend to act in ways that disadvantage them with regard to knowledge distribution.

What about conceptual or interpretive resources? Above, I called these some of the raw materials for knowledge—that's why they're part of the epistemic economy. The concepts available to a thinker structure her thought; lacking the concept *blue*, Homer called the sea 'wine-dark' and said oxen were likewise 'wine-looking'.⁸ The poet had many timeless thoughts; apparently "oxen and oceans are different colors" was not among them. But what we can't think, we can't know. And of course this is the point of Fricker's invocation of the Carmita Wood case: Wood can't know she's been sexually harassed if she has no such concept.⁹

In what follows, I'll understand hermeneutical resources somewhat more broadly than this, though: it's not just about what concepts we have; it's also about how they fit together. One way to put the point is that concept possession involves not only competence in naming—calling all and only the Fs, 'F's—but also competence in *responding* to *F*-ness, recognizing what

8. British Prime Minister William Gladstone was the first to highlight Homer's strange color talk; Deutscher (2011) has much more on the development of color concepts (and words) from the ancient world to the present.

9. As will be more apparent later, I think the Carmita Wood case, as commonly discussed, oversimplifies considerably.

follows—in terms of inferences, interactions with other applicable concepts, etc. So even if I was perfectly accurate at labeling all the ducks in the world, I wouldn't be fully competent with the concept *duck* if I didn't grasp how duckness relates to concepts like *egg* and *swim*, and how duckness is incompatible with truckness, and so on.

And of course, as soon as we're talking about concepts that have to do with items that are not just natural kinds, things get complicated fast. To adapt an example from Haslanger (ms, 2), to be competent with the concept *cooking* involves grasping human roles—*preparer* and *consumer* of food—which can be filled by the same person or by different people; and also a whole host of other concepts, from *food* and *heat* to *cuisine*—and those concepts are inescapably tied to complex systems of farming and distribution, labor and culture, and so on. In other words, concept possession ends up invoking a bunch of *practices*: far more than a question of naming, it's about scripts and narratives and relations (between people, between things, and between people and things) and social meanings, a complex web of culture.¹⁰ All of which is to say that concept possession involves a host of dispositions, both to identify *F*-ness and to respond to it, where the appropriate responses often require following culturally-given scripts and where the connections between concepts will be captured in a web of narratives. In short, as the title of Haslanger's paper suggests, 'cognition [is] a social skill.'

This is a big, complex area, and I can't do it justice here. But what we need for the purposes of this paper is just to recognize that when I say that one class of epistemic goods is *hermeneutical resources*, I'm not just talking about words. Hermeneutics, *interpretation*, requires words *plus* the web of their connections, which is structured by schemas and narratives. Knowing

10. I follow Haslanger (ms, 3) in taking practices to be "patterns of behavior that enable us to coordinate due to learned skills and locally transmitted information, in response to resources, and whose performances are 'mutually accountable' by reference to some shared schemas/social meanings." What I say here is intended to be entirely compatible with her view on this.

you've been sexually harassed isn't just a matter of being disposed to utter those words as a label for a way you've been treated; it's about recognizing what it means, and what it calls for, to have been treated in that way. This was, of course, why sexual harassment was called 'sexual harassment' in the first place—the coiners of the term were consciously appropriating the legal concept of *harassment*, importing the set of narratives and scripts about harassment's wrongness, systems of redress, and so on.

So: I'm saying that there is a complex system of production, reproduction, and distribution for a class of epistemic goods, i.e., knowledge and other epistemic resources, two kinds of which are *credibility* and *conceptual/interpretive resources*. What can we say about this system? And how does thinking about it help us understand epistemic injustice?

The epistemic economy is a social structure, and it is embedded in and interacts with other social structures large and small. For example, the epistemic economy involves the educational system, which is a set of structures and practices that function (in part) to create and disseminate knowledge and mark those who possess it;¹¹ in this capacity the epistemic economy interacts with the 'regular' economy, which takes advantage of the output of the educational system—both the knowledge and the credentialing—in the labor market, the intellectual property market, etc. At a more intimate scale, families also interact with the epistemic economy, both in that there are practices of knowledge production and distribution that operate within families and in that the familial practices are both influenced by and inputs to the larger system. Parents teach children, in the segment of the epistemic economy that operates within the family, and parents are generally taken to be authoritative with regard to what's best for children, in the extra-familial epistemic economy. In short, the epistemic economy is complex, massive,

11. I say 'in part' because in the present-day United States the educational system has other functions as well, such as providing the site of an enormous and lucrative sports industry.

and entangled with pretty much every other major social structure.

Because the epistemic economy is a social institution, it is a potential site of injustice. There can be injustice in the distribution of epistemic goods, for example, and the goods that are unjustly distributed can be of instrumental or intrinsic value. There can also be injustice that is not a matter of distribution. In what follows I'll try to make clear which kind(s) of injustice are operative in particular cases, but the kinds I'll be discussing will usually be familiar. If you know what distributive justice is in other domains—say, health care, or the job market—you're equipped to understand it here; it's a different good being distributed, but the injustice in distribution will be of an ordinary kind, for example.

Given this reorientation, I'll take a moment to say a bit more about how I propose we understand testimonial and hermeneutical injustice in turn. I think that with a few small adjustments we can translate Fricker's individualistic conceptions to something suited to my more structural approach.¹²

3.1.2 TESTIMONIAL INJUSTICE, REVISITED

If my proposal is that we take epistemic injustice to be the general category of injustice that occurs in and through an epistemic economy, how am I suggesting we think about testimonial injustice? I propose two main changes, which will also support a terminological adjustment.

First, unlike Fricker, I'm saying that testimonial injustice is to be understood as the *pattern* of reductions in credibility suffered by members of marginalized groups. It thus continues to exist, on my way of thinking, even at a moment when (or if), somehow, no particular speaker happens to be being disbelieved on the basis of her membership in a historically marginal-

12. I should say that given the sprawl and complexity of the epistemic economy, it's possible and even likely that there are other forms of epistemic injustice beyond testimonial and hermeneutic, but I won't investigate that question here.

ized group. I think understanding testimonial injustice as a pattern makes better sense of the way those who navigate it have to incorporate it into their thoughts and plans, as a feature always present in the world they inhabit. I'll say more about this below, but first I want to say a bit about another way I'm going to diverge from Fricker in thinking about testimonial injustice.

Specifically, I think the phenomenon should be understood to reach well beyond testimony. I agree with Fricker that what is at issue is reductions in credibility on the basis of social identity, but I think that testimony is only one of many places that the problem manifests. Yes, it is unjust if members of marginalized groups are treated as less credible in their assertions—and it is also unjust, and unjust *in the same way* as in the testimonial cases, if members of marginalized groups are treated as less credible (i.e., less well-positioned epistemically, and/or less likely to behave in a way appropriate to their epistemic position) in other situations.¹³

What I mean is that it is not only unjust that campus police officers very often disbelieve what young Black men tell them about whether they are students at the universities they attend. It is also unjust, and in the same way, that women are very often treated as if they don't know what they are doing, even if they're not talking at all. One woefully common manifestation of this is of course *mansplaining*, as when very skilled women at the climbing gym are barraged with unsolicited, unnecessary, and unwanted 'helpful pointers' from men who are much less capable.

So as I'll be using the term, testimonial injustice refers to a structural feature of the epistemic economy, in which there is a pattern of members of historically marginalized groups being subject to reduced credibility across

13. Recall that *credibility*, as I understand it here, is similar to the epistemic *trustworthiness* discussed in Daukas 2006 (111), which says "A is epistemically trustworthy if and only if A is disposed to behave (when contextually appropriate) as though her epistemic status is S if and only if her epistemic status is S."

a wide range of circumstances, including but not exclusive to testimony. Because I want to include cases that are not about testimony, I will refer to this kind of injustice as *credal injustice*.

Credal injustice is a clear distributive injustice. It's distributive because it has to do with the allocation of credibility in the epistemic economy; it's an injustice because the allocation of credibility is not based on people's actual epistemic positions (or behavior relative to those positions) but rather on their social identities as members of marginalized groups. Credibility is clearly of enormous instrumental value: I have defined it as a prerequisite for full participation in the epistemic economy. But I don't want this instrumentality to mask the depth of the harm that credal injustice can involve: the ability to be and to conceive oneself as a knower depends on being granted the access and recognition that credibility supports. Which is to say that there's an intrinsic harm here, along with the instrumental: epistemic agents subjected to credal injustice are significantly cut off from their epistemic selves. As Fricker says, they are harmed specifically in their capacities as knowers.

I'll say more about the many faces of credal injustice below, and I'll work through more cases. But for the moment, what we need is to register that credal injustice, as I'm thinking of it, differs from Fricker's conception in those two ways. I think it's structural, and I think it's about more than just testimony. Epistemic injustice in the form of credal injustice can exist even if it is not being instantiated in a particular transaction at a given moment, because it's a matter of patterns in the social world. And epistemic injustice in the form of credal injustice occurs whenever marginalized people are treated as less credible just because of their social identities; mansplaining is an injustice of the same kind as Fricker's cases of failure to believe some bit of testimony.

3.1.3 HERMENEUTICAL INJUSTICE, REVISITED

As with testimonial/credal injustice, I want to diverge from Fricker in understanding hermeneutical injustice as a structural feature of epistemic economies,

and then I also want to make a second adjustment.

Since hermeneutical injustice, even on Fricker's view, is about the presence or absence of communal interpretive resources, it just seems natural to think of it as a structural feature of the epistemic economy, having to do with patterns in what is produced by the system. So this is a place where Fricker's insistently individualistic methodology seems especially odd. But aside from that, we might again ask if she is arbitrarily limiting the range of cases in ways we might want to avoid.

Consider the famous case of the missing concept *sexual harassment*. It's natural to think that something has gone wrong if a community's conceptual and interpretive resources are simply silent on large swaths of significant experiences, just because those experiences belong to marginalized people, even aside from questions about particular transactions in which failures of intelligibility occur. As Stanczyk (2012) argues, justice concerns not only how we choose to distribute the things we produce, but what we choose to produce in the first place. Taking hermeneutical injustice to be a structural problem, over and above its operation in particular transactions, takes this thought seriously while still allowing productive discussion of the kinds of particular transactions that concern Fricker.

Furthermore, beyond the structural versus transactional question, I want to ask: Why take hermeneutic injustice to be solely a matter of *missing* interpretive resources? Why not think that sometimes the issue will be about *defective* interpretive resources, or conceptual tools that are imperfectly grasped? Just as I wanted to agree with Fricker that reductions in credibility were the issue in testimonial/credal injustice while broadening the range of cases in which we recognize those reductions as occurring, so here I want to agree with Fricker that reductions in intelligibility are the issue in hermeneutical injustice, while recognizing a broader range of causes for those reductions.

Once again, recall Fricker's paradigm case: the missing concept of *sexual harassment*. It's not, to me, quite obvious that the lack of that concept

was the (only) problem; I'm inclined to say that part of the problem was that the experiences of people being sexually harassed were understood to be experiences of, say, *flirting*, where the application of an existing concept that suggested benign-ness and even desirability, rather than some other existing concept (perhaps something like *lewdness*) deepens the wrong. In other words, the problem isn't that Carmita Wood and others had no words, that they suffered from a pure absence of conceptual tools to understand what was going on; it was that the community provided tools that affirmatively misrepresented what was going on.¹⁴

In short, even this supposed paradigm case looks rather more complex to me than Fricker's discussion suggests. And I think that in fact there are a number of ways the state of a community's interpretive resources and patterns in grasp of those resources can reduce intelligibility in ways that are harmful and unjust to members of marginalized groups. I'll describe some more examples in detail below; for now, I hope that my suggestion that even the Carmita Wood case has more facets than Fricker supposes will give you something of the flavor of the revision I propose.

So just as with credal injustice, I want to understand hermeneutical injustice as a structural phenomenon, and I want to broaden the range of cases in which we recognize its occurrence. Just as I think we should focus on the

14. There's are some distinctions worth making here. Treating what happened to Wood as an instance of *flirting* could be willful or not; the concept *flirting* might be defective or not. In one scenario, 'flirting' does not capture what Wood experienced, but someone calling what she experienced 'flirting' just doesn't fully grasp what flirting is. In this case, the misapplication is not willful, and the concept is not defective; we've just got operator error in the use of a concept. In another scenario, 'flirting,' as used in the community, really does include what Wood experienced. But this would (one could argue) mean that *flirting* was in some sense a defective concept. In yet another scenario, we can imagine that calling what Wood experienced 'flirting' is a willful misuse of the concept, that will be, if systematically undertaken, similar to certain propagandistic uses discussed in Stanley 2015 (50ff and throughout). As he notes (see 50ff), the language of liberal democratic ideals is often deliberately (mis-)used, coopted, to cover up the gaps between those ideals and reality. In what follows, I'll try to be clear about which of these possibilities is at work in particular cases.

idea of patterns of reductions in credibility and not limit ourselves to cases of testimony in particular, so also I want to focus on the idea of reductions in intelligibility, and not limit ourselves to cases only of *absent* interpretive resources. If we go my way, we can see how there are both distributive and non-distributive injustices here. And we can see that there are both instrumental and intrinsic harms caused by this form of injustice as well: hermeneutical injustice causes instrumental harm when it prevents Carmita Wood's getting the redress she deserves; it causes intrinsic harm when it puts knowledge about important aspects of marginalized people's identities out of reach (about which more below).

3.1.4 THE KEY DIFFERENCES IN MY PICTURE COMPARED TO FRICKER'S

Fricker gets a lot right. But I think she misunderstands where epistemic injustice is really located, and I think she misses some important kinds.

To summarize: Unlike Fricker, I conceive of epistemic injustice as fundamentally a structural feature of epistemic economies, rather than a property of particular individual interactions. This affects how I understand both credal and hermeneutical injustice, shifting the focus to patterns rather than individual transactions.

I also make a further amendment in conceiving each of these forms of epistemic injustice, broadening their application. In the case of credal injustice, I think the relevant pattern is found not only with regard to testimony, but also with regard to assessments of members of marginalized groups' credibility *in general*. In the case of hermeneutical injustice, I think the relevant issue is not just about the lack of needed interpretive and/or conceptual resources, but also certain kinds of defective use of such resources, whether because the resources are not fully developed, or because they are defective, or because the agent has only partial grasp of or access to them. Fricker misses all these additional kinds of epistemic injustice; I think they're important.

Given all that, here are amended definitions of our central terms:

Epistemic injustice (according to me): structural injustice in the operation of the epistemic economy, which can include, possibly among other forms:

Credal injustice (my version of testimonial injustice): a pattern of unjust reductions in credibility for members of historically marginalized groups¹⁵

Hermeneutical injustice (my version): a pattern of unjust reductions in the intelligibility (either to herself or others) of facts relating to people who are members of marginalized groups, due to defects in the community's interpretive resources or members' grasp and/or deployment thereof

A note: The inclusion in these definitions of a specific reference to marginalized groups may seem surprising, but I won't argue for it here. I'm following Fricker, who thinks that while individuals may be wronged when they (for example) are given less than the appropriate amount of credibility in some exchange for reasons other than prejudice connected to social identity, that is not *injustice*. I find this thought even more plausible when we shift the location of the fundamental problem to the structural level, and I suspect that the kinds of patterns I'm leaning on are both evidence and instances of oppression of the kind that is constitutive of marginalization. But that's all I'll say about the question here.

As promised, I think some advantages of my amendments to Fricker's picture should now be clear. 'Going structural' allows several of our initial questions to get answers of a kind we're familiar with. We can assimilate questions about what kind of injustice this is to more familiar general types of distributive and non-distributive injustice. We need not locate the injus-

15. It seems plausible to me that there are also patterns of unjust *increases* in credibility for members of historically dominant groups, and these also count as credal injustice in my sense, but in this paper I'll focus on the reductions.

tice solely in particular individual transactions; structural features are the fundamental issue.

As a result, several of our opening questions now seem less pressing: we need not theorize about how a hearer can wrong a speaker by failing to believe her, because we can understand the problem as one of distribution, instead of grappling with the puzzles about doxastic voluntarism and other thorny issues that naturally accompany the former project. A similar move allows us to think about whether and how medieval serfs could be personally wronged by their lack of Marxist theory: there really is a question of justice about what kinds of interpretive resources (and thus knowledge) a society produces and does not produce, which is distinct from claiming that Marx's precise work was unjustly withheld before he turned up to bestow it.

On the question of instrumental versus intrinsic harms, and how the harms of epistemic injustice are distinctive, my picture gives us reason to expect not only instrumental harms having to do with the distribution of knowledge, but also further, distinctive harms flowing from the particular operation of our two flavors of injustice and the workings of the epistemic economy. I've said a bit about some of these further harms, and I'll continue to try to delineate the precise contours of harm as I discuss additional particular cases below. There will also be more to say on this front when we've got the other piece of the puzzle—my view of intentional action—on the table, so that we're considering the interaction between epistemic injustice and action as I understand it.

So let's talk about action.

3.2 WHAT IS INTENTIONAL ACTION?

If you were watching me type this paper, you might ask, "Why are you scowling at the screen?" And I might give one of two kinds of answer. On one hand, I might say, "I didn't know I was doing that!" On the other hand, I might say, "It's how I psych myself up to tackle an objection. Duh."

Anscombe (1963, 9ff) says ‘why?’ questions like these—questions where the sense of ‘why?’ is “that in which the answer, if positive, gives a reason for acting”—mark off the class of intentional actions. When I not only have no answer as to why, but deny any knowledge of doing the thing asked about, my scowling is unintentional, a mere side effect of an overdose of philosophical effort. But in the other case, when I am aware of the scowling and offer a reason, my scowling is intentional: it is then not a side effect but a prescription to counteract dangerously low levels of philosophical ferocity. When you act intentionally, Anscombe said, you *know what you’re doing*, and how and why.¹⁶

And this seems reasonable, at first blush: it seems right to think that if I have no idea I’m doing something, I can’t be doing it intentionally. The problem is that it’s just not true that we always know what we’re doing when we act intentionally; there are lots of counterexamples. The most famous comes from Davidson (2001, 91–92), who points out that we often undertake things while in grave doubt about whether we’ll succeed—and that doubt can be so substantial that we do not even believe we’re doing what we intend, even though, if we pull it off, we clearly will have done so intentionally.

So, for example, suppose I am scowling at the screen as I type because I need to increase my argumentative zeal, and I’m *hoping* that the scowling will do the trick. I may type for quite some time, diligently scowling, fighting through an increasingly achy brow, without believing that it’s really going to help. But all that perseverance means that any increase in ferocity that the scowling actually produces will certainly not be an accident—it will be intentional, despite all my doubt.

16. She adds that you know it in a special, direct way—‘without observation’ or inference—but I won’t make use of that specialness in my arguments here. See Anscombe 1963, 13–15, 49–50 for her discussion of the non-observational, non-inferential nature of this knowledge; see my (ms a) for how I understand its specialness and distinctive practicality.

So Anscombe has offered a tempting thought, but there's a serious problem. I propose we avoid the counterexamples by making a key revision.¹⁷ *Practical knowledge* is Anscombe's name for the special agential knowledge of what one is doing (and how, and why) that enables answers to 'why?' questions. She takes the possession of practical knowledge to be a necessary condition for intentional action—but the counterexamples make that untenable. Where has Anscombe gone wrong? I argue she's right that there's an important relationship between intentional action and practical knowledge; her mistake is characterizing it as one of necessity. Instead, I show, practical knowledge is not a necessary condition but a *normative condition*: agents possess it *when things go well*.

Which is to say that intentional actions *aim at* instantiating practical knowledge. When, and to the extent that, an agent fails to know what she is doing, how, and why (in the special way), her action is *defective as such*. So if I am ϕ -ing, but in doubt, so that I don't know I'm ϕ -ing, things are not going well, or not as well as one would hope, from the point of view of intentional action.

The benefit of my adjustment is that an agent who is in Davidsonian doubt can still be acting intentionally even in the absence of (full) practical knowledge—that's the point of moving from a necessary condition to a normative one. Using the normative condition, we get the intuitively right result: when my doubtful scowling does its job, that outcome *is* intentional; it's just that things weren't going as well for me, qua agent, as they might have. (After all, I'd prefer to know either that I'm doing what I intend, if only to avoid the stress of wondering, or that I'm not, so that I can abandon the attempt.)

I can't provide the full argument for the view here; that project is the work of my (ms a). But here are some key points.

17. The full argument for my view is found in my (ms a).

If we're interested in what intentional action is, we need to distinguish it from several nearby phenomena. Intentional actions are different from idle predictions, wishes that come true, failed attempts and lucky successes, cases of deviant causation, etc. And I argue that what distinguishes intentional actions is precisely that they, and only they, are defective when and to the extent that the agent fails to have practical knowledge. A wish is a wish whether the agent knows it is coming true or not—indeed, we can wish for things in rightful certainty that they are not coming to pass—and so on. The connection between intentional action and an aim of practical knowledge, I argue, is about the distinctive kind of agential *control* that intentional action is supposed to involve.

I make my case by reflecting on skillful actions, which involve an elevated level of practical knowledge compared to less-skilled versions: when Serena Williams plays tennis, she knows an enormous amount about what she's doing, how, and why; when I play, pretty much everything that happens is both a surprise and a mystery. Serena's knowledge is control; my lack of knowledge is my lack of control. Serena's tennis actions are worlds better, qua actions, than mine, because she has all that knowledge, i.e., control, and I don't.

This means that intentions, on my view, must involve belief. An intention contains a proposition about what one will do: I will ϕ by means m for reason(s) r . When things go well, as in Serena's tennis playing, this belief is knowledge, because it is accurate and not as a matter of luck. When I play tennis, even if I, say, hit a winner, things aren't going well in the Serena way, in the way that's about practical knowledge, because it *is* a matter of luck. When things *really* go well, according to the standard of intentional action, what happens is both what the intention describes, *and because* the intention describes it—the intention guides the action and secures the outcome. That's what practical knowledge is: an *executed*, not just fulfilled, intention. And that's control.

On this way of thinking, it's clear that just as control comes by degrees, so

does practical knowledge. And that means that to say intentional actions are distinguished by an aim of control—i.e., practical knowledge—is to say they are subject to a *graded* standard; they are defective *when and to the extent* that the agent fails to achieve that aim.

For the purposes of this paper, it's important to recognize that agents can fall short of full practical knowledge—and so be acting defectively—in several ways.

I've already mentioned one: the agent might doubt that she is doing what she intends, because she is unsure about her ability (to, say, scowl her way to success). But there are other ways. For starters, if intentions are instances of the schema *I will ϕ by means m for reason(s) r* , then in order to know the whole complex (as practical knowledge requires) an agent has to know a variety of other propositions. She has to know things about her means, including that her means are such that, if she takes them, she will ϕ . She has to know the facts that serve as her reasons.

Of course, the agent might have an attitude to some of those propositions about her means or her reasons that falls short of full belief, and thus holds her short of full practical knowledge in the same way that the doubt about her ability does in our initial counterexample. Furthermore, she might know (or believe to some degree) these facts in more or less detail—and the extra detail makes for more control. So the gradedness of control can manifest in different degrees of confidence about propositions involved in the intention (about her circumstances, her means, her ability, etc.), or in knowledge of more or fewer of those propositions.

Finally, recall our discussion of hermeneutical resources. An agent can have better or worse grasp of the concepts involved in her intention—in all the

ways discussed in §3.1¹⁸—and that variation in level of grasp can affect her level of practical knowledge. If you're trying to achieve *F-ness*, but you only dimly grasp what *F-ness* involves, it will be a matter of luck, rather than the control that practical knowledge provides, if you achieve it.

To summarize: on my view, the distinguishing feature of intentional action is an aim of control, and that control is Anscombean *practical knowledge*—the special knowledge agents can have of what they're doing, how, and why. When an agent acts intentionally, things go well when and to the extent that she has control, i.e., practical knowledge. In the good case, her intention describes what happens, and not as a matter of luck: it describes what happens because it *shapes* what happens.

I capture my way of thinking in a principle:

NEED TO KNOW: When you act intentionally, you aim at control, i.e., practical knowledge; when (and to the degree that) you fail to have such control/knowledge, your action is *defective*, though it can still be intentional.

3.3 SHACKLING

Suppose (as I'll ask you to do for the rest of this paper) that I'm right about intentional action, and intentional actions are defective as such when and to the extent that the agent fails to know what she is doing (and how and why) in a special way, as NEED TO KNOW says. Then, when there is epistemic injustice in the sense I described in §3.1, we should expect certain interactions between the nature of action and the operative epistemic injustice. Why? Because epistemic injustice affects what people (can) know, and how people are assessed as knowers—and knowing things is relevant to

18. Remember that I think concepts are part of a web of culture that also includes narratives, scripts, and so on. So grasping concepts requires grasping that whole rich apparatus.

meeting the standard of NEED TO KNOW, and being recognized as knowing things is relevant to being recognized as meeting the standard of NEED TO KNOW.

In other words, if I'm right about action (as I'm asking you to suppose), and if there's epistemic injustice in the world (as there is), then we should expect at least two kinds of problems, one for each of our two kinds of epistemic injustice. First, hermeneutical injustice will occlude facts having to do with marginalized agents in a way that will make it difficult or impossible for those agents to know what they're doing, how, and why as NEED TO KNOW requires. Second, credal injustice will cause observers to assess marginalized agents as failing to know what they're doing, how, and why as NEED TO KNOW requires.

But that's not all. For example, we should also expect more complex versions of the two main types, including combinations. So in the next section, I'll discuss a series of real-world examples and discuss how they instantiate the phenomena I'm predicting here. But I do want to take a moment to discuss what these formal characterizations tell us.

The key thought is that the combination of NEED TO KNOW and epistemic injustice will tend to put members of historically marginalized groups in a double bind: they must either act defectively, given their epistemically unjust circumstances, or not act at all, in deference to those defects. The combination of epistemic injustice and NEED TO KNOW limits their choice to wrongdoing or inaction. This is the dilemma I call *shackling*.

The possibility of shackling is easiest to see in a case where hermeneutical injustice makes the agent unable to know some fact that is, for example, serving as her reason for ϕ -ing. Carmita Wood couldn't know that her reason for quitting her job was because she'd been sexually harassed, because hermeneutical injustice made that fact unavailable to her. But because she couldn't know that fact, she couldn't act on it with practical knowledge—because practical knowledge requires knowing what you're doing, how, and why. So her only options were to act defectively, or not to act at all, in the

face of her circumstances.

Credal injustice makes things more complicated: it's about how agents are assessed with regard to their epistemic positions, not those epistemic positions themselves. So as a first pass, we should expect a slightly different kind of shackling in credal injustice cases: the choice will be between taking an action that will be *assessed as* defective (because the agent will be assessed as failing to know what she's doing, how, and why) and not acting.¹⁹ This is the plight of the women at the climbing gym mentioned above: they know that they must either go climbing and face mansplaining, or skip the gym. Act in a way that will be wrongly assessed, or forego acting.

There are a few points I want to flag before we get to our cases.

One is about what kind of harm shackling causes to marginalized agents. As we'll see, shackling can make it harder for an agent to achieve the things she wants and deserves—for example, it often interferes with her ability to know her means, and that's an important instrumental harm. But it's more than that.

When I say that epistemic injustice is an obstacle to action that is non-defective by the standard of NEED TO KNOW, it's important that NEED TO KNOW is the *constitutive standard* of intentional action. Making it impossible for an agent to satisfy NEED TO KNOW is a harm to her *in her capacity as an agent*. To see this, notice that hermeneutical injustice affects not just an agent's means, but also the kinds of reasons that are available for (non-defective) action. And the kinds of reasons that hermeneutical injustice makes it harder or impossible to access are very often things that are central to the identities and experiences of marginalized people. (Again, think of Carmita Wood.) But if you can't act from the deepest parts of your identity, or your most important experiences, there is a way in which your

19. Though as we'll see below, these assessments can be self-fulfilling, so that the agent will be forced to choose between *actually* defective action and inaction after all.

actions can't be truly yours. This barrier to acting from central features of one's identity—and so making the actions you take properly your own—is thus a deep, intrinsic harm, separate from any instrumental problems, and it's distinctive of the picture I'm articulating here. Shackling alienates marginalized agents from their actions.

Another question one might have involves the way that characterization of shackling offered here suggests a moralized reading of the defectiveness of shackled actions. If the defects are epistemic, what's this talk of *wrongdoing* about? As I hope the cases below will make clearer, I think there are good reasons to accept that there will in fact tend to be a moral tinge to assessments of these defects.

First, since these defects are (actual or erroneously assessed) failures to act well *as such*, there's a sense in which they are precisely (actual or erroneously assessed) failures to act as one ought. Which is to say, they *are*, in a broad sense, (actual or erroneously assessed) moral failures.

Second, as we'll see in a type of complex case discussed below, sometimes the kinds of narratives involved in grasping concepts about the social world will invoke the moral side of the dual nature of the trust required for epistemic cooperation noted by Daukas (as discussed on p. 91 above)—imputing lack of sincerity, or benevolence, etc. to members of certain groups—as a way of justifying the reductions in credibility that characterize credal injustice: so both aspects of Daukasian trustworthiness will be in question.

Finally, because in a social world our actions just do affect others, failure to act well by the standard of NEED TO KNOW will often be seen to be a failure to take due care with regard to other people. None of that is to say that the defects at issue in shackling are always, in a narrow sense, moral. But it's why I'm not concerned if you hear a somewhat finger-wagging note in the definition of shackling. It's apt, given shackling's status as a practical, wide-sense moral, problem, and the intermingling of epistemic and (narrow-sense) moral features in many pieces of this puzzle.

One final point before we look at some cases. Stanley (2015, especially

chapter 6) considers the relationship between epistemic injustice (Fricker’s version) and a claim linking knowledge and proper action, namely the claim that knowledge is interest-relative; I want to distinguish my view from the one he presents. Stanley endorses the claim that knowledge is interest-relative, which is to say he thinks that two agents, identically epistemically positioned with regard to some proposition, may differ in whether they know that proposition just because of differences in what is practically at stake for each of them with regard to its truth.²⁰ (So—to borrow a case often used in this literature—if you and I are each considering whether to stand in a long line at the bank on a Friday afternoon in order to deposit a check, or to wait until Saturday instead, and we have exactly the same evidence that the bank will be open on Saturday, there can be a difference in whether each of us *knows* the bank will be open if only one of us is in danger of bouncing the rent check if we don’t get the deposit in on time.) Stanley goes on to argue that this means that marginalized people will be unable to know facts about their oppression, precisely because the stakes are too high for actions that implicate those facts.

But in her (2018) reply, Dotson brings out some deep problems with Stanley’s argument. Stanley cites thinkers (including Dotson herself, as well as, e.g., Wright and DuBois) who are members of precisely the groups that Stanley says will be unable to know facts about oppression; as Dotson notes, Stanley’s citations of these thinkers as authorities is incompatible with the claim he is arguing for. Furthermore, as Srinivasan discusses at length in her (2016) critique of Stanley, the claim that marginalized people will be less able to know facts about their own oppression than more privileged people is in conflict with a long tradition of compelling Black, feminist, and Marxist thought that says that marginalized people are the *best* positioned,

20. In my (ms b) I argue that the evidence used to support claims of interest-relativity—a collection of intuitions about cases—does not in fact support that claim, and suggest that my view better captures the phenomena while being supported by a ground up argument (i.e., my ms a). But I won’t rehash that discussion here.

the *most* likely to know facts about oppression—that marginalized people often possess *double consciousness*, because they see the world both as their oppressors see it (as they must, since the oppressors are in control and what the oppressor sees will affect what happens) and as it is (since only by understanding the actual, oppressive systems can they navigate them).²¹

But my view does not have the problems identified for Stanley by Dotson and Srinivasan. For one thing, as I discuss at length in my (ms b), my view about the link between knowledge and non-defective action does not entail interest-relativity; indeed, it can be used to debunk it. Moreover, while I am indeed suggesting that there will be cases in which hermeneutical injustice creates obstacles to marginalized people’s knowledge of important facts, including facts about their oppression, nothing in my account suggests that this is always, or inescapably, the case. We can see this by noticing that Stanley’s view looks incompatible with the possibility of consciousness-raising, since attempts at consciousness-raising do not change the stakes of action against oppression; by contrast, my view strongly suggests consciousness-raising as a necessary remedy, because it is a way of creating a new community that can incubate new and better hermeneutical resources. (It’s not a coincidence that the conceptual work on *sexual harassment* occurred through just such a process.) I do not have space to go into great depth on this subject, but I hope these brief thoughts will suffice to distinguish my claim from Stanley’s, and to stave off Dotson- and/or Srinivasan-esque worries.

So, to summarize before we (at last) take up some cases: We should expect at least two key types of interaction between epistemic injustice and the nature of intentional action as captured by NEED TO KNOW.

When hermeneutical injustice meets NEED TO KNOW: By making it difficult or impossible to know certain facts, hermeneutical injustice will make it difficult or impossible to take certain actions (namely

21. Stanley replies to his critics in his (2018).

actions involving those facts) non-defectively.

When credal injustice meets NEED TO KNOW: By making it difficult or impossible to be assessed as knowing various facts, credal injustice will make it difficult or impossible to be assessed as taking certain actions (namely, actions involving those facts) non-defectively.

Recognizing these interactions allows us to identify *shackling*, a dilemma that can be both an instance of and a means to oppression.

Shackling: Agents who are members of marginalized groups will often be forced to choose between acting defectively (or acting in a way that will be *assessed* as defective) and not acting, because epistemic injustice makes it too difficult to know (or be assessed as knowing) facts relevant to what those agents are doing, as non-defective action requires.

I've said that we can expect to see not only clean cases of a single flavor of shackling, but instances where both the "direct" and the assessment-based kinds of shackling are at work in a single case, instances where unjust assessments are self-fulfilling, instances of interactions between the two kinds. But we'll get to all that.

That's as far as I think we can go in the abstract. Let's look at some cases.

3.4 CASES

In the last section, I gave an abstract characterization of the basic kinds of interactions we should expect to see between circumstances of epistemic injustice and the nature of intentional action. A key phenomenon arising from that interaction is the one I call *shackling*: To be shackled in my sense is to be forced, by the interaction of epistemic injustice and NEED TO KNOW, either to act defectively (or be assessed as acting defectively) or not to act. It is an agential dilemma. But we might wonder: what do real-world instances of the schemas I presented look like? How common are they? How serious?

To start answering those questions, we'll need to look at some cases. So that's the project of this section.

3.4.1 SHACKLING VIA HERMENEUTICAL INJUSTICE

Hermeneutical injustice, as I've been saying, is a pattern of unjust reductions in intelligibility of facts relating to (members of) marginalized groups, due to certain kinds of defects in the community's interpretive resources or the community members' grasp and/or deployment thereof. In other words, this is the circumstance of missing, defective, or ill-used conceptual tools, where the patterns of absence, flaws, and/or misuse align with patterns of marginalization. And as I described in the previous section, the paradigm case of shackling via hermeneutical injustice occurs when facts relating to marginalized people are obscured in these ways, thereby making it difficult or impossible for agents from marginalized groups to act non-defectively by the lights of *NEED TO KNOW*.

In order to act non-defectively, *NEED TO KNOW* says, agents must have practical knowledge; having practical knowledge requires knowing facts about one's circumstances, and one's means, and one's reasons—any fact that figures in one's intention. Since hermeneutical injustice makes many facts relating to marginalized people less available, actions that depend on those facts will often be defective. Or anyway, that's the idea. So what are some examples?

Several things I've said so far have hinted that we can read the Carmita Wood case as one such example. As Fricker describes her, she lacks the concept of sexual harassment, as does her interlocutor, so her experience is unintelligible to both of them. On this reading, the failure of intelligibility, both to Wood and to her interlocutor, is the hermeneutical injustice, and it's a matter of a missing concept, a gap in the community's interpretive resources. But I read the case somewhat differently.

Remember that I made two adjustments to the Frickerian conception of hermeneutical injustice. First, I said that we should understand it as a

structural phenomenon, not a matter of individual interpersonal transactions. So where Fricker thinks the injustice only exists when a particular failure of intelligibility occurs, I think it persists as long as the community's interpretive resources don't do justice to the lives and experiences of marginalized people. A failure of intelligibility with regard to some fact entails that there is a fact that could be intelligible; I think it is an injustice that some facts—facts having to do with marginalized groups—are less intelligible in virtue of the way our interpretive resources happen to function whether or not any actual failures of intelligibility are occurring. After all, for an attempt at intelligibility to occur, someone has to realize that there is something that might be intelligible. But if the available resources are inadequate enough, that realization may be impossible. To my eye, such a case is not an absence of hermeneutical injustice, but an extreme case.²²

Which brings me to the second way I differ with Fricker about hermeneutical injustice. Remember, I said that lacking important concepts is not the only way a community's hermeneutics can be unjust. Hermeneutical injustice also occurs, on my view, when important concepts are imperfectly grasped—either because they are obscured by other features of the epistemic environment or because they are defective such that they could never be fully grasped (as they would be if, say, they were internally contradictory). When we have imperfect grasp of a concept, we tend to err in a few ways: we may take things to be *F*s that are not *F*s; we may fail to recognize *F*s as *F*s; and we may respond inappropriately to *F-ness* (failing to make appropriate inferences, take appropriate actions, and so on).²³ I think that this

22. Consider an analogous case of absent conceptual resources where that absence is not a question of justice. Before the advent and eventual widespread acceptance of the germ theory of disease, humanity was unable to prevent or remedy some of its greatest afflictions. Things were clearly not going well just because no one was trying and failing to understand what microorganisms were.

23. I am neutral on questions the ontology of concepts, but I am committed to the idea that conceptual *grasp* is a matter of having certain dispositions: dispositions both to correctly identify and to correctly respond to *F-ness*.

expansion helps us get a better grip on the Carmita Wood case, and also see just how common cases of this form are.

Let's look at the Wood case through this lens. As I said above, it seems likely that Wood herself, and potentially others with whom she discussed her experience, might well have thought of what happened to her as *flirting*. We can understand that possibility in one of two ways: perhaps calling what happened to Wood *flirting* was apt, in which case there was something defective about the public concept—I'll say more about defective concepts below—or perhaps calling what happened to Wood *flirting* wasn't apt, in which case community members were just not fully competent with the concept of flirting. (Because if they had been fully competent, they would have realized that touching and talk of the kind she endured are not flirting if they are repeated and unwelcome advances made in a workplace under conditions of significant power asymmetry, for example.)

So it's not just that the right label wasn't available. It's that a wrong label was. And the application of the wrong label suggested all kinds of responses that were not the appropriate ones to what actually happened: if it was flirting, perhaps she should have been flattered. Perhaps she had no reason to quit. And even if Wood wasn't, herself, tempted to think that what happened was flirting—even if she was inclined to think it was more like abuse, or sex discrimination—the fact that the patterns of concept application in her community would emphasize *flirting* and not abuse or sex discrimination will mean she will find it hard or impossible to make herself understood. (And of course, the resistance she'll face along the way may make her doubt her own insight.)

So that's how the Wood case looks different on when read as an instance of my version of hermeneutic injustice. I think this reading does better at capturing the complexities of the case, while still clearly being of a piece with the 'missing concept' thought that motivates Fricker. But there's one more step we need: how is the Wood case an instance of *shackling*? How does this epistemic injustice become agential injustice?

The answer, of course, is that it will be shackling if the issues with partial concept grasp make it impossible for Wood to have practical knowledge in acting. Which is exactly what we see.

Notice that Wood is not just experiencing confusion about what happened to her; she is trying to *make herself understood*, and she is trying to *respond to what happened*. She is trying to *do things*, to act on the basis of her experience. But she can't do that, because without the concept of sexual harassment, she can't know she's been sexually harassed—so she can't, say, have practical knowledge that she's quitting her job *because she's been sexually harassed*, since she can't know her reason.

And that's not all. Suppose she intends to quit her job because she's been abused, and/or because she's experiencing sex discrimination. I take it that it's true she was abused, and true that what she experienced was sex discrimination; to say that she experienced sexual harassment is to sharpen those claims, not to deny them.

So suppose Wood intends to act for the reason that she's been abused, or that she's been subjected to sex discrimination. I'm stipulating that she can know she's been abused, and/or subjected to sex discrimination. But it's natural to think that the right response to abuse is to call the police, and that the right response to sex discrimination is to sue. Given the patterns of concept application in relevantly similar cases, the police would do nothing and if a lawyer took her case she would lose—remember, the general thought in the community is that what happened to her was *flirting*. So she's likely to fail in her attempt to seek redress, because of failures of uptake on the part of her audience.

But even that is not yet the crux of the problem, from the point of view of NEED TO KNOW. The real heart of the matter is that even if, miraculously, Wood did find a lawyer who took on her case (maybe just for publicity, maybe out of real solidarity), and, even more miraculously, managed to win a judgment (again, for reasons good or bad)—well, that would be so miraculous, such a stroke of luck, that it would not have been a case of

practical knowledge. It would have been far too out of Wood's control. If she intended to get compensation for what was done to her, she couldn't know she was doing that—the system was not structured to make legal redress an outcome after sexual harassment. The concepts and narratives and scripts weren't set up that way.

Finally, remember that I emphasize that hermeneutical injustice is structural, not transactional. If Wood had somehow gotten lucky, she still wouldn't have succeeded in acting with practical knowledge, but the very fact that it was a matter of luck shows that the structural problem exists for agential injustice just as much as for epistemic injustice. Women *like* Wood—who were subjected to sexual harassment—are, as a class, subject to shackling. The pattern need not hold with perfect universality to exist and to matter for justice.

So Carmita Wood faced a dilemma. She could take action in response to the fact that she was sexually harassed, but that action would necessarily be defective. Or she could forego acting in response to that experience. Both options are bad ones: for there to be no way for her to fully and properly exercise her agency in light of a profound experience is a deep harm, one that goes beyond the financial and social costs she incurred. That's the double bind of shackling. Wood couldn't know she'd been sexually harassed, but she knew for sure that what had happened to her was something she had to respond to. But because of her community's interpretive resources, there was no way for her avoid being shackled.

I've talked through the Wood case at such length for two reasons. First, because it's familiar from the literature on epistemic injustice, and, second, because it offers a way to bring out features we can find in additional cases. So let's look at some of those features in isolation to see what other cases they bring to light.

The first feature to consider is the way that imperfect grasp of relevant concepts can be an obstacle to non-defective action. Let's think about how that works in a little more detail, and then a little more generally.

I think we might want to say that Wood had an extremely early stage of grasp of the concept of sexual harassment. She had an inchoate sense that her experience was of a kind that had certain features and merited certain responses—it seemed bad, and it seemed like it merited quitting her job, and not with a penalty. She could in some sense point to *this thing* that happened, and think ‘*That* is a reason to quit, and I’ll tell human resources that so they can help me.’ She had some of the dispositions that go with grasping the concept of sexual harassment: dispositions to associate certain kinds of treatment in the workplace with to-be-avoided-ness and to-be-rectified-ness, etc.; perhaps also the sense that that treatment related to phenomena like *abuse* and *sex discrimination*. (Those ill-defined dispositions are of course the raw materials with which the people who went on to develop the concept of sexual harassment began their work.) But the inchoateness of Wood’s grasp of what happened to her meant not only that she lacked the word for it, but also that she couldn’t know what responses were called for. The abuse-like qualities suggest police; the sex-discrimination-like qualities suggest civil action; the lack of any established conceptual structure along with the bare fact that this happened in the workplace suggested trying HR. Her minimal grasp of the concept meant that she had no access to what follows from its aptness, and so no ability to act in response to it.

And that’s how conceptual grasp affects agents’ ability to have practical knowledge, in general. Here’s the picture. Conceptual grasp is graded, running from total lack of grasp to perfection. Increases in grasp are increases in what can be known and what can be done. Let’s track Sarah as she comes to grasp the concept *salsify*.

1. Sarah is on the subway and overhears part of a conversation. One passenger says, “So what did he say?” The other says, “Salsify.” (*She is now disposed to acknowledge that ‘salsify’ refers to something in English, though this will be a very weak disposition since it could be a foreign word and in any case she has no idea if it’s a noun or a name or an adjective or a verb...*)

2. That night, Sarah is at a restaurant. On the board where the specials are posted, she reads “Crushed celeriac and roasted salsify \$12.” *She is now disposed to take salsify to be a food item.*
3. Sarah asks the server what salsify is. He replies that it’s a root vegetable, sort of like a parsnip. *She is now disposed to associate it with fall and winter, things appropriate for vegetarians, etc.*
4. Sarah orders and eats the celeriac and salsify special. *She is now disposed to recognize roasted salsify by sight, and to associate the flavor of salsify with the taste of artichoke hearts, though this is again a weak disposition since she’s not sure how to distinguish what part of the dish’s flavor comes from the celeriac and what part from the salsify.*
5. Some time later, Sarah is at the farmer’s market and sees a bin of brown things that look like sticks with a sign that says ‘salsify.’ *She is now disposed to recognize it by sight when raw.*
6. Sarah gets a new job, working for a chef who loves salsify and grows it on the property, cooking it in and participating in the development of many dishes. *Sarah is now disposed to recognize the taste of salsify alone and in combination, cooked and raw, to expect it to taste good with certain things and bad with others, to take it out of the oven after x minutes, and so on.*
7. Sarah takes over the cultivation of the salsify in the restaurant garden. *She is now disposed to harvest it at a certain point, to water it under certain conditions, and so on.*
8. Sarah studies with a botanist and learns about salsify’s relation to other plants. *She is now disposed to recognize it as a relative of the dandelion.*
9. Sarah studies with a biologist and learns about salsify’s role in wild ecosystems. *She is now disposed to identify it as a food item for certain animals and not others, and so on.*

10. Sarah studies with a historian and learns about salsify's popularity in the Victorian era, and that it was one of Thomas Jefferson's favorite foods. *She is now disposed to look for it if she tours the grounds of Monticello, to bring it to a friend's period potluck, and to recognize it under its alternate name, 'Johnny-go-to-bed-at-noon.'*

At step 1, it looks like Sarah has the most minimal possible grasp, or perhaps a precursor for the most minimal grasp, of the concept *salsify*. At step 2, I would say she has the concept, though her grasp is definitely very basic. Her grasp increases as we progress through the steps. By step 4 or so it seems right to say she is competent in an ordinary sense with the concept; by step 6 or 7 she seems to have expert grasp of the concept; if we keep going in the vein of steps 8–10, maybe adding some time with a shaman to learn about traditional medicinal uses and some time with a chemist to learn about basic structure, we'll be approaching perfect grasp of the concept, where that would mean being disposed to associate salsify with all its features, and to draw all the correct inferences and take all the correct actions with regard to salsify—grasping all the nuances of salsify's place in the conceptual web.

Once we lay out the progress of Sarah's grasp of salsify stepwise like this, it's easy to see how increasing grasp of the concept increases the reach of Sarah's agency vis à vis salsify. If she's at step 4, and intends to get some salsify by foraging, it will be a matter of luck if she manages to do so—she can't have practical knowledge that she's getting some salsify by foraging if her grasp of what salsify is doesn't extend to recognizing it in the wild. Her action will be defective, because she won't have control. Where as the Sarah of steps 6, 7, or 8 *can* non-defectively get some salsify by foraging.

In short, what the Sarah case shows is how increasing grasp of a concept can increase what an agent can know in ways that are relevant to her action: the better she grasps a concept, the more she can do—the more actions she can take while having control, i.e., practical knowledge. Increasing conceptual grasp increases what can be known, and so what can be done non-defectively.

As Sarah learns more about *salsify*, she is able to take the fact that *this is salsify* as a reason to do more things—at step 4, she’s able to take the fact that *this is salsify* as a reason to expect an entree at another restaurant to be delicious; at step 6 she’s able to take the fact that *this is salsify* as a reason to substitute it for asparagus when the kitchen runs out of the latter.

So as Sarah’s grasp of the concept increases, she calls more and more of the bits of salsify in the world ‘salsify,’ and she calls fewer and fewer of the non-salsify things in the world ‘salsify,’ and she draws the salsify-appropriate inferences and takes the salsify-appropriate actions in a wider and wider range of circumstances. She knows more propositions involving salsify, and I’m inclined to say that there are some propositions involving salsify she knows *better*, precisely in virtue of the way that her knowledge of something like *salsify is a root vegetable kind of like a parsnip* could mean that she can distinguish it from numbers and dollars and doughnuts as well as root vegetables not like parsnips—say, round ones like turnips—at step 3, but her knowledge of that same proposition at step 7 and beyond could mean that she can make all the step 3 distinctions and more, say because she also grasps the ways salsify is not like a parsnip, and what salsify’s being like a parsnip means for how it tastes and roasts and grows, and so on.

I hope this digression has made clear how conceptual grasp relates to knowledge and action.²⁴ I hope it’s also apparent how this same kind of progres-

24. It’s worth noting that there are two potential claims I might be making about the way merely partial grasp of a concept can be a barrier to knowledge (and so to non-defective action). One claim would be that if I have only partial grasp of *F-ness*, then I can’t know any proposition involving that concept. This would be an alarmingly strong claim. But I don’t need anything that strong. What I need is something much weaker. For my purposes, we can say that Sarah knows *that’s salsify* when she sees some at the farmer’s market after step 5, even though she has certainly not achieved the perfect grasp of salsifyness that she might achieve at or beyond step 10. The reason we can say she knows it’s salsify when she sees it is that her grasp of salsifyness includes what it looks like raw. On the other hand, step 5 Sarah’s grasp of salsifyness *doesn’t* extend to recognizing it under a microscope or by smell, or to associating it with dandelions

sion is possible with terms that are not about natural kinds. We'll want one more thing before we use all this, plus the Carmita Wood case, to generate more cases.

The last thing we need is to understand how the epistemic economy can cause problems with conceptual grasp that constitute hermeneutical injustice of the kind that leads to shackling. The short answer is that concepts relevant to the kinds of facts for which hermeneutical injustice comes into play are sometimes prone to being imperfectly grasped through *neglect*, and sometimes through *concealment*. I'm not going to say much about concealment here. Concealment cases will often involve Stanley-style propaganda, as mentioned above; another kind of concealment will occur when, for example, sex ed courses are designed to be silent on everything other than heterosexual cisgendered monogamous relationships. But let's talk about neglect.

When neglect is at work, conceptual resources will tend to be *underdeveloped*—as if the community investigated *salsify* to, say, step 3 and no further. For socially significant concepts, neglect frequently results from taking the dominant group to be the default. So, for example, women and members of racial and ethnic minorities have often been unable to know what is happening and respond appropriately when they have certain health issues, because all the research and education has focused on how the issues in question manifest in white men. Take the hermeneutics of *heart attack*, for example. The grasp of this concept both among the general public and within the medical profession have involved descriptions of symptoms that are paradigmatic in men but not at all in women—so women have been less likely to realize them-

and Thomas Jefferson. So what I'm saying is that not just the level of grasp but the particular contours of that grasp will matter for what Sarah can know. The aspects of salsifyness at work in the proposition and/or the action must match the aspects of salsifyness Sarah grasps. (These thoughts are similar in flavor to the view advocated in Stanley (1999), on which the truth of an ascription of understanding is context-relative. But I don't have space in this paper to explore the relationship between that view and mine.)

selves that they were having heart attacks, and so to seek treatment, and to be recognized by treatment providers as having heart attacks even when they do seek treatment. The community's partial grasp of *heart attack* is a hermeneutical injustice, and it leads to shackling: women are unlikely to be able to act non-defectively for the reason that they're having heart attacks because that fact is obscured by the available interpretive resources. As in the Wood case, it will be a matter of luck if they hit upon the appropriate diagnosis and response.

Treating the dominant group as the default isn't the only way that these problems can arise, of course. Each year, several thousand children from China and Korea are adopted by families in the United States. Many of the adoptive families are white. Children adopted into these white families often at some point wish to explore and express the identities and cultures of their birth. But in the U.S., the distinctions between Chinese identity and culture and Korean identity and culture and, indeed, Japanese and Mongolian and Indonesian identity and culture are not well captured by the available interpretive resources; American culture tends to lump all these and more together into an undifferentiated *Asianness*.²⁵ So a child born in China but raised in America is unlikely to grasp what Chineseness is well enough to express it in her actions non-defectively. If she managed to do something authentically of Chinese culture, it would likely be a matter of luck: it's not clear that her environment offers her the resources to grasp that learning aikido, for example, would not be a way of connecting with her roots, whereas making niangao might be.²⁶

25. It is not only Americans who have this problem. In July 2018 the British Foreign Secretary, visiting China, accidentally referred to his Chinese-born wife as Japanese in a meeting with his Chinese counterpart. His gaffe was widely diagnosed as arising from his habitually thinking of her as *Asian* rather than Chinese.

26. Aikido is a Japanese martial art; niangao is a traditional Chinese New Year cake, most popular in Eastern China.

So we now have a recipe for one way hermeneutical injustice can lead to shackling: the hermeneutical environment may make it difficult or impossible for agents from marginalized groups to have adequate grasp of concepts relating to their experiences (as in a Carmita Wood or woman's heart attack case) and/or identities (as in the case of the adopted child). These agents won't be able to know facts about their reasons and their means that they must know in order to act non-defectively: facts like *I was sexually harassed* or *I'm having a heart attack* or *this is a way of expressing my birth identity*. This puts limits on the ways these agents can exercise their agency—limits that bar non-defective action in domains that are critically important both instrumentally and non-instrumentally.

I want to flag here that only my view can make this clear. Why? Because my view's requirement of *knowledge* for non-defective action makes the difference. If you had an alternate view, on which action doesn't require knowledge, but rather just belief, or a 'distinctively practical attitude' as adherents of Bratman's (1987) way of thinking suppose, then these cases will look just fine, even though in the original Carmita Wood case, where we understand it to be a matter of a completely absent concept, we'll see a problem. That's because in the original Wood case, the complete absence of a concept means that Wood can't even have a *thought* about sexual harassment—can't believe, can't have a distinctively practical attitude, none of that. So clearly she can't act on any of those attitudes. But in the partial grasp cases the agents can and do have thoughts—that's why they're trying to act from those thoughts. Partial grasp of a concept doesn't prevent thought, or belief. It's the knowledge requirement that brings out the way things are going wrong for the agents in the partial grasp cases—the knowledge requirement that is distinctive of NEED TO KNOW.

So hermeneutical injustice—an unjust deficit in available interpretive resources—makes it the case that partial-grasp Carmita Wood, and our women heart attack victims, and our adopted Chinese and Korean kids are unable to fully know propositions that are necessary for their intentions (and actions) to be non-defective. They are shackled: they can act defectively on the basis

of their incompletely-grasped experiences of sexual harassment, or medical conditions, or birth identities, or they can refrain from acting on those things—but as I’ve been stressing, inaction on facts like these is hardly a viable alternative. That’s the daily, damaging dilemma.

And notice that on this view, something has gone wrong even if our agents avoid disaster. Even if our heart attack victim survives, and our Chinese born child ends up learning to make niangao just like her birth family’s, it won’t be *because of* what they intended; it won’t be because what happened was under their control in the way Serena Williams’ tennis playing is under her control. That’s the deep harm of shackling, and it’s only visible on my view: the way epistemic injustice limits the reach of marginalized people’s agency, removing the possibility of control.

Of course, something like this is what Stanley (2015) was trying to get at in his discussion of interest-relativity. But the problem with Stanley’s view was that he thought the issue was about what’s at stake for marginalized people, which didn’t leave room for double consciousness and the ways that the marginalized both can and do often have *extra* knowledge, knowledge that can enable resistance. On Stanley’s view, acting to resist will destroy knowledge, because of the way that action will raise the already high stakes for the marginalized. But as I said above, on my view, we can see how the community of the marginalized can be a source of new and better concepts, concepts that can be the means to resistance, and the stakes don’t have any bearing on knowledge.²⁷ So we can see both how shackling via hermeneutical injustice is an unjust and substantial burden while recognizing the possibility and necessity of change. Consciousness-raising and other improvements to

27. I should say that I think stakes can *indirectly* affect knowledge, because it seems right to say that *belief* is interest-relative: high stakes can make people doubt in ways that are belief-undermining, and so, by extension, knowledge undermining. But that kind of interest-relativity doesn’t have the upshot Stanley claims: it’s not necessarily the case that knowledge is destroyed by high stakes—the very community solidarity that can enable consciousness-raising can be a source of confidence that counteracts the tendency of stakes to produce doubt.

the available conceptual resources can and must happen.

3.4.2 SHACKLING VIA HERMENEUTICAL INJUSTICE, PART 2

The partial grasp cases above show how distributive injustice in hermeneutical resources leads to shackling. But recall that on my view hermeneutical resources are rich and complex, having to do not just with single concepts like *sexual harassment* or *heart attack* or *Chineseness*, but also narratives, and scripts, and schemas. And defects in those narratives and scripts and schemas due to hermeneutical injustice can likewise lead to shackling.

So imagine a young Black man, call him Michael, involved in a traffic stop. He knows how dangerous that situation is for people like him, so he intends to be respectful, compliant, and in no way suspicious. He calls the officer “sir.” He meets his eye, to avoid seeming shifty, but doesn’t hold it, to avoid seeming defiant. He reaches for his license and registration when asked, but moves very slowly, to avoid seeming as if he’s going for a weapon. Even if Michael takes all this care, because being *respectful* and not being *suspicious* depend in part on uptake, he can’t know he’s doing what he intends because the hermeneutical patterns of police officers are capricious. With some officers, on some days, Michael’s actions will be read in the way he intends; with other officers, or on other days, his amount of eye contact will be read as falling too far on the shifty side (or the defiant side), or the speed with which he reaches for his license will be read as too slow, and so reluctant, or sassy, or too fast, and so dangerous. He can’t know if the officer will read his behavior through a ‘thug’ lens or not. So even if he gets lucky, and the officer interprets him as respectful, compliant, and not suspicious, Michael can’t know that he’s doing what he intends, because the hermeneutical environment infects the officer’s responses with caprice.

This is a case where it’s natural to think that Michael has double consciousness—he has a lot of knowledge about the defectiveness of the prevailing hermeneutical resources, and is making every effort to navigate them in light of that knowledge. But he can only do so much: it’s not just Michael’s reading that matters; he’s at the mercy of the officer’s interpretation. The dominant cul-

tural narratives prescribe respectfulness and forbid acting suspicious. But narratives around blackness make it unclear if or how that's possible, as the tensions in the description of Michael's choices bring out. The means to Michael's end are obscured by hermeneutical injustice, because of his dependence on the officer's response. For him, the shackling is double-layered: hermeneutical injustice (on the officer's end) makes it the case that Michael can't non-defectively take action to safely navigate the traffic stop; he can't know what he's doing in the way NEED TO KNOW requires, because of the caprice in the officer's response. But in Michael's case, the other horn of the shackling dilemma isn't quite available, either: if Michael tries not to take the defective action, the caprice in the officer's response means he will probably be interpreted as taking some (bad) action anyway. This kind of conversion of the inaction horn into a second-order defective-action horn will be a hallmark of these audience-interpretation cases.

This case is similar to our earlier partial-grasp cases in that the issues arise from problems with the availability and deployment of hermeneutical resources. One key difference is the more pronounced role of the audience's interpretation as compared to the agent in the Michael case. Another is that while there are some concepts that play roles in Michael's case like the role played by *heart attack* or *Chineseness* in those cases—*respectfulness*, for example—the Michael case leans more on the additional components of a hermeneutical environment. In the Michael case, it's complexities like the interactions between (e.g.) the script for being respectful and the narratives around blackness that generate the caprice that puts the outcome beyond Michael's control.

Before moving on, I should reiterate that this is, once again, a case where a mundane yet important situation cannot be navigated via the marginalized person's agency. So even if Michael comes through unscathed, things have not gone well for him. The lack of control he experiences is an oppressive harm *whether or not* he suffers physical or other harm. It is also a case where only my view allows us to see this: it's the knowledge requirement that captures the control problem, just as in the partial grasp cases above.

And of course, what Michael's case really is is a manifestation of a structural injustice: young black men, as a group, face this problem; even if sometimes things don't end in disaster, the pattern is one of undue limitations on agency.

3.4.3 SHACKLING VIA CREDAL INJUSTICE

What about shackling that happens through credal injustice? Recall that credal injustice, as I'm using the term, is about patterns of unjustly reduced credibility. Combined with NEED TO KNOW, credal injustice means that agents from marginalized groups will be *assessed* as failing to know (or know fully) the facts about what they're doing, even when they know perfectly well. And since NEED TO KNOW is not just an internal standard—it also operates explicitly in our evaluations of actions—actions involving the facts for which the agents are unjustly assessed as not (fully) knowledgeable are therefore unjustly assessed as defective.²⁸ So in these cases, the shackling dilemma is not a choice between acting defectively and not acting; it's a choice between acting *such that one will be assessed as doing so defectively* and not acting. All that was predicted in §3.3. So where do we see it happening?

To get a feel for what these cases are like, we can again imagine a slightly altered version of the Carmita Wood story, using some thoughts introduced above. Imagine Wood being interviewed, during the process of determining her eligibility for benefits, about her departure from the job where she was sexually harassed. Now vary the case just a little and suppose that when she is asked why she left the job, instead of saying she doesn't know, she

28. That we routinely criticize agents for failing to know facts relevant to their actions is the basis of a huge swath of literature on contextualism and interest-relativity in epistemology; see Hawthorne & Stanley (2008) for a representative example. The restaurant case with which they begin has become standard: In that case, two people have a time-sensitive restaurant reservation. When one of them takes them down a side street on the way there, she is criticized for doing so without knowing the restaurant is there.

says, “my boss abused me.” Her interviewer asks her to describe the abusive behavior; when she does, the interviewer says “Come on—you don’t know that he was being abusive; it sounds like he was just flirting with you!” As in the real case, she is seen as having left her job for no good reason, and while it’s easy to see that there’s a dispute about applicable concepts here, it’s also true that the problem in the dispute is about an unjust lack of credibility: Carmita is treated as not knowledgeable about her own experience. To put it in the terms we saw in the schema above, our alternate Carmita is wrongly assessed as not knowing what she’s doing (leaving her job because she was abused), because she’s wrongly assessed as not knowing the reason for which she acts (the fact that she was abused).²⁹

This is common for marginalized people; there are many current, actual cases with the same structure. For example, in a number of jurisdictions, women seeking abortions are prevented from getting one unless they undergo an ultrasound (sometimes an invasive transvaginal ultrasound) and watch the output of the procedure as a doctor describes their fetus. To the extent that promoters of these laws are acting in good faith, this is a case of these women being assessed as not knowing what abortion is, or what a fetus is, or some similar fact necessary for them to know what they’re doing, despite abortion’s being one of the most-discussed medical procedures in our society and despite the very basic bodily function implicated. As a result they are barred from acting until they are “informed.”

The instinct to discount women’s knowledge in this way is so ingrained that many people accept the idea that these forced ultrasounds serve to inform women, rather than manipulating them. But note that we do not have similar requirements in any other similar cases. A father can consent to high-risk surgery for his child without being told he must observe the gruesome procedure beforehand or speak to a parent whose child did not survive the procedure. A man can have a vasectomy without having to watch *Three*

29. Just as in the restaurant case mentioned in the previous note.

Men and a Baby. Potential organ donors do not have their driver's licenses withheld until they visit a trauma center to see for themselves the horrible cost of the shortage of donors. Only abortion is treated this way, because the real issue is that women are not treated as knowledgeable and authoritative about their own reproductive lives. They are shackled: these laws assess women seeking abortions as lacking necessary knowledge and thus as acting defectively, and require that they therefore either forego the abortion or be "informed" by the ultrasound.

In another case, women in media circulated among themselves a Google spreadsheet containing the names of men in media who have engaged in objectionable sexual behavior, intending to warn others. An outcry ensued because "you don't know what happened." Once again, the claim is that the women don't know the proposition that is serving as the reason for their action, and so that the action is defective. (There is an implied additional claim that women who, upon encountering the list, take actions like *avoiding the men named* are also acting defectively, because they too, we are to think, fail to know their reasons.)

Because of the role outside assessment plays in this form of shackling, in many of these cases we see people forced to choose the inaction horn of the shackling dilemma: they are prevented from acting due to wrongful assessments of their epistemic position vis a vis propositions that figure in their intentions. This happens even when in fact their intended actions are non-defective; I take it women do know what an abortion is, what a fetus is, etc, when they seek abortions.

Notice, too, that this kind of wrongful assessment can sometimes be self-fulfilling: sometimes the manipulation of the forced ultrasound will convince women that they were confused about what they were doing, even though they were not; sometimes the insistence that women don't understand their own sexual experiences makes them doubt their knowledge in that domain. When that happens, we see the similarities between this type of shackling and the Michael case, in which the interpretive flaws in his audience pre-

vented his knowing what he was doing.

So while there are differences more and less subtle between the kinds of cases we've considered thus far, the similarities suggest we should look for interactions. Call those *combined cases*.

3.4.4 A COMPLICATION: COMBINED CASES

Combined cases are cases in which both hermeneutical and credal injustice are interacting with NEED TO KNOW. One key domain where we see this is (again) women's sexuality. The combined epistemic injustice that operates here even has a name: *rape culture*. Rape culture is the set of narratives that ratifies and enforces the reduced credibility ascribed to women in sexual contexts. Patterns of reduced credibility for women (and others who don't identify, or are not identified as, men)—credal injustice—are explicitly endorsed and justified by the hermeneutical framework, which is used to support patterns of action assessment and response. So the credal injustice is part of a larger unjust hermeneutical environment.

This interaction manifests in cases. For example, given the narratives that say that women are ignorant and/or dishonest about their sexual experiences, women seeking to report rape, or warn the community about sexual harassers, can't know that they're doing so—because the speech acts of reporting and warning require uptake, and that uptake is withheld because of the narratives. So even if these women have evidence, and they speak up, and they take all the steps in all the ways that, in other circumstances, are adequate to the task of reporting or warning, they can't satisfy the requirements of NEED TO KNOW. (As is often noted, reporting a mugging requires much less of the victim than reporting a rape.) So rape culture serves to shackle women: they must either refrain from reporting, or do so defectively.

So there are actions (here, speech acts) that marginalized people can't know they're performing, because of the *interaction* of credal and hermeneutical injustice. Unjust narratives in the hermeneutical framework generate and support reductions in credibility. Those reductions in credibility create toxic

doubt in the agent, causing her to lose her grip on relevant concepts (e.g., *rape*) and so lose knowledge of the very thing she set out to act on. The negative assessment of her is self-fulfilling, hermeneutical injustice creates credal injustice creates hermeneutical injustice, and the victim is shackled twice over.

3.4.5 A FINAL COMPLICATION: DEFECTIVE CONCEPTS

Earlier I mentioned that problems with the application of hermeneutical resources can have any of several causes. Sometimes agents are learning a concept, and don't yet have complete competence with it. Sometimes people aren't really still learning, but just have less than full grasp—the way, say, someone who slept through most of high school physics and never thought about it again might still have a grasp of *gravity* that's not all that much better than Wile E. Coyote's. Sometimes concepts will be willfully misapplied, perhaps in an effort to reshape a hermeneutical environment, as in the propaganda cases that are of central concern in Stanley (2015).

I want here to flag a special kind of case: *necessary* misapplication. Necessary misapplication will happen when concepts are inherently defective—so that the problem isn't with the agent's grasp of a perfectly good concept; rather, the problem is that the agent could never fully grasp the concept, because taken in total it is contradictory. There is a long tradition of critical thought that suggests that the ideologies of capitalism, and patriarchy, and white supremacy all possess deep but hidden contradictions.³⁰ I haven't relied on this idea in what I've said so far—so if you don't like it, you don't have to revise your opinion of the previous material—but I want to mention what it means for the picture I'm painting if we *do* accept it.

Suppose that all, or nearly all, our concepts around gender, race, capital, etc. are inherently defective. I find this very plausible: for one thing, it would

30. Classic loci for this thought include Marx (2004) and Beauvoir (1953); it's affirmed on both sides of the debate about Stanley (2015): he and Srinivasan (2016) agree.

explain the prevalence of double binds, because we'd see that our conceptual frameworks contain contradictory narratives, incompatible scripts and mandates about what to do, and so on. (It's easy to see how the Michael case could be read in this light, and it's not hard to think of additional cases, either.) If we have whole swaths of concepts that are profoundly broken, then whole swaths of *everyone's* actions will be defective according to NEED TO KNOW. Is that a problem for my account?

I don't think so. In fact, I'm inclined to think both that it's right, and that it's just more grist for my mill. Because we can see that even if everyone is in the same boat with regard to these concepts, everyone does not pay the same price for the problems they create. Shackling cases involving these concepts may be all but universal, but assessments of agents' actions are not on a par. Members of dominant groups are assessed as acting non-defectively even as members of marginalized people are criticized and worse. And we've seen above that injustices in assessments of action are themselves important forms of shackling.

One final point about defective concepts. We might wonder if this is a case where a Bratmanian gets the same result I do, or if I am still offering something unique. The question is whether there can be a thought, enough material for a belief, or a distinctively practical attitude, when an inherently defective concept is involved. If not, then Bratman can say as much as I about necessary misapplication cases. So for my account to be the only way to capture this problem, it must be that we can still have thoughts, beliefs, and 'practical attitudes' involving inherently defective concepts; the thing that's out of reach has to be *knowledge*. And I think that's right.

To see why, start by recalling how Sarah's grasp of *salsify* progressed above. In the beginning, she could make a few distinctions, know a few very local facts. As her grasp grew, so did her epistemic and agential reach.

Now consider a slur. I'm going to make one up, to avoid using a real one. Suppose 'flark' is a derogatory term for people born in Connecticut, and suppose there's a history of oppression of these Connecticut-born people

(as ill-grounded as actual oppressions are), so that this is a case like real slurs. Someone learning the slur might first come to understand it via the non-derogatory ‘born in Connecticut’ content. And it seems to me that at that stage, this slur-learner can have thoughts when she is applying *flark*, even though there is no proposition ‘x is a flark’ that is true, since that proposition would require that there be something wrong with being born in Connecticut. The slur-learner’s use of *flark* does pick out something in the world, just not what she supposes. (It picks out people born in Connecticut.) The slur-learner will never fully grasp or fully competently use *flark*, because that’s not possible. She could never know ‘x is a flark,’ because ‘x is a flark’ is always false. But she can have a false belief that ‘x is a flark,’ and she can intend to meet a flark in a Bratmanian way, and so on.

I think the slur-learner and her mismarked flarks show us how one can have thoughts and beliefs and practical attitudes involving inherently defective concepts, even as it is impossible to have knowledge and act non-defectively while using them. When a white female police officer tells Black schoolchildren to be respectful if they are stopped by police, she’s having a thought there: she’s invoking a respectability script that would work, in a limited range of cases—ones where the person trying to be respectful is, say, a white woman like her. It’s just that that script *only* works locally; in a context where it’s interacting with blackness, the behaviors mandated by the respectability script don’t work anymore, as we saw in the Michael case.

All of which is to say that while there’s much more to explore on this subject, I think we can see that inherently defective concepts fit well with the view I’m advocating, even though I did not lean on them in my initial presentation.

3.5 CONCLUSION

This is a very big paper, so let’s take stock.

I agree with Fricker that epistemic injustice is about unjust harms to knowers *qua* knowers—it’s about barriers to the possession of knowledge, and

barriers to recognition as a bearer of knowledge. But I began by introducing a conception of epistemic injustice that's different from Fricker's classic presentation in some key details. First, I take epistemic injustice to be, fundamentally, a structural phenomenon, rather than something that exists only in particular individual transactions. Second, I think that both of the two kinds of epistemic injustice Fricker identifies—what she calls testimonial and hermeneutical injustice—are broader than she supposes. Testimonial injustice isn't just about testimony, on my view (so I propose a terminological change to *credal* injustice); it's about reductions in marginalized people's credibility more generally. Hermeneutical injustice, on my view, isn't just about gaps in collective conceptual resources; it's about a range of injustices in the community's production, reproduction, and use of its interpretive tools.

With that understanding of epistemic injustice in hand, I sketched my view of the nature of intentional action. I summarized my view in the principle NEED TO KNOW, which says that actions are defective, *qua* actions, when and to the extent that the agent fails to know how to have *practical knowledge*. On my view, actions are only fully successful when an agent knows, in a special, direct way, what she is doing, how, and why—because this knowledge is the special kind of control that is the whole *point* of intentional action: this control is intentional action's *constitutive aim*.

But if epistemic injustice creates barriers to possessing and/or being recognized as possessing knowledge, and non-defective action *requires* knowledge, then we should expect some bad interactions when marginalized agents are acting, or trying to act, in circumstances of epistemic injustice. I laid out the interactions we should expect in abstract terms, and then I walked through some cases to flesh things out. What we saw was the phenomenon I call *shackling*: the combination of NEED TO KNOW and epistemic injustice will tend to put members of historically marginalized groups in a double bind—they must either act defectively, given their epistemically unjust circumstances, or not act at all, in deference to those defects. The combination of epistemic injustice and NEED TO KNOW limits their choice to wrongdoing or

inaction. It eliminated the possibility of agential control.

We saw that this creates problems of both distributive and non-distributive justice, and that the harms are both instrumental (since shackling makes it harder for marginalized agents to get what they want and deserve) and intrinsic (since shackling harms marginalized agents *qua agents* when it limits the reach of their agency and the ability to act in ways necessary for their actions to be fully their own). Shackling thus manifests and constitutes oppression. And given that shackling commonly operates in domains relevant to *resisting* oppression, shackling also reinforces and protects the prevailing order.

My view of action is the only avenue to understanding all this, as we saw. Popular alternative views about the nature of action, because they don't link proper action to *knowledge*, can't explain some of the crucial phenomena, such as the way that something has gone wrong, oppression is still present, even when outcomes are not at their worst. Our Michael case has something of the flavor of Philando Castile's story, but I have shown how Michael need not suffer Castile's fate to be harmed. Oppression is not only present in physical violence; the picture I've laid out here shows its subtler, more insidious aspects.

One way to see the importance of this point about shackling—that it shows how something is going wrong even when things in some sense go better than we might expect—is to consider a related point often made by political philosophers who worry about domination (like Pettit 1997) and feminists as far back as Astell and Wollstonecraft. The thought is that living under a dictator is a harm even if that dictator is benevolent; just being subject to arbitrary rule is a bad state of affairs. My view brings out the way that culture can be such a dictator. Even when it smiles upon Michael, so that he is able to leave the traffic stop unscathed rather than dead, he is harmed just in virtue of being subject to its caprice.

So one larger upshot of this paper is to support an old claim of feminist thinkers: justice cannot be secured either just by changing individual at-

titudes or just by changing the laws and institutions of the state.³¹ *Culture*—understood as the web of practices and their associated hermeneutical resources—is a both a site and a source of injustice, and that has to be reckoned with.

There is more work to do, even aside from that project. I suggested that there may well be additional forms of epistemic injustice that we will find via an effort to create a more exhaustive list of epistemic goods; these new forms of epistemic injustice are likely to yield new forms of shackling. And of course, there's more to say about the kinds identified here.

As I confessed above, I'm sympathetic to the idea that shackling affects every one of us, due to the deep contradictions hidden in some of our central conceptual structures—but either way it is clear that it affects the most vulnerable among us most powerfully of all. I've tried, here, to shed light on these chains. I hope that can be one step toward breaking them.

31. See Haslanger 2017 for a forceful articulation of this idea.

REFERENCES

- Anderson, E. (1999) "What is the point of equality?" *Ethics* 109(2): 287–337.
- . (2012) "Epistemic justice as a virtue of social institutions." *Social Epistemology* 26(2): 163–173.
- Anscombe, G.E.M. (1963) *Intention*. Cambridge, MA: Harvard University Press.
- Bratman, M. (1987) *Intention, plans, and practical reason*. Cambridge, MA: Harvard University Press.
- Daukas, N. (2006) "Epistemic trust and social location." *Episteme: a journal of social epistemology* 3(1–2): 109–124.
- Davidson, D. (2001) *Essays on actions and events*. New York: Oxford University Press.
- Deutscher, G. (2011) *Through the language glass: why the world looks different in other languages*. New York: Picador.
- Dotson, K. (2018) "Distinguishing knowledge possession and knowledge attribution: the difference metaphilosophy makes." *Philosophy and Phenomenological Research* 96(2): 475–482.
- Fricker, M. (2007) *Epistemic injustice: power and the ethics of knowing*. New York: Oxford University Press.
- . (2013) "Epistemic justice as a condition of political freedom?" *Synthese* 190: 1317–1332.
- Haslanger, S. (2017) "I–Culture and critique." *Aristotelian Society Supplementary Volume* 91(1): 149–173.
- . (ms) "Cognition as a social skill."
- Hawthorne, J. & Stanley, J. (2008) "Knowledge and action." *The Journal of Philosophy* 105(10): 571–590.
- Jaques, A.E. (ms a) "Where there's a will, there's a way of knowing."
- . (ms b) "Knowing what you're doing with shifty epistemology."
- Pettit, P. (1997) *Republicanism: a theory of freedom and government*. Oxford: Oxford University Press.

- Srinivasan, A. (2016) "Philosophy and ideology." *Theoria: Revista de Teoría, Historia y Fundamentos de la Ciencia* 31(3): 371–380.
- Stanczyk, L. (2012) "Productive justice." *Philosophy & Public Affairs* 40(2): 144–164.
- Stanley, J. (1999) "Understanding, context-relativity, and the description theory." *Analysis* 59(1): 14–18.
- . (2015) *How propaganda works*. Princeton, NJ: Princeton University Press.
- . (2018) "Replies." *Philosophy and Phenomenological Research* 96(2): 497–511.
- Williamson, T. (2000) *Knowledge and its limits*. Oxford: Oxford University Press.

4 KNOWING AND DOING, WELL AND GOOD

Hannibal Lecter is good at stuff. He's good at manipulating people; he's good at giving the cops the slip; he's good at carrying on a conversation with his victims even as he slices open their skulls and nibbles on the still-busy brains inside.

The guy really knows what he's doing. He's got everything under control. Hannibal Lecter looks like a really, really successful agent.

I mean, obviously he's also super evil. But so what? Wouldn't anyone be better off if they could just be as good at doing stuff as Hannibal Lecter?

This paper takes that question seriously. It offers a view of the nature of intentional action on which control is the peak of agential achievement. Even more than that, it's a view on which the *only authoritative standard* for action is this aim of control. And then it asks: if that's really what it takes to act well, what's wrong with being a cannibalistic criminal mastermind?

Here's the (non-diabolical, I swear) plan. In §4.1, I'll sketch my view of the nature of intentional action. In §4.2, I'll show how that view of action entails a new form of constitutivism about practical reason; then, in §4.3, I'll suggest that if what I've said is right, we may be living in a world where Hannibal Lecter is pretty much an ideal agent—because it's not clear what could be defective about his evil ways. In §§4.4–4.6, I'll sketch some possible avenues of escape from Lecter's clutches; then I'll briefly address

some lingering questions in §4.7 before concluding.

This paper is a bit unusual. You'll see that it's all based on my view of the nature of intentional action, which I describe but don't argue for.¹ Instead, I just ask you to suppose that I'm right about action, so that we can explore what it would mean. Then I raise a big worry about what follows from my view—the specter of Lecter—and I offer only a partial and tentative response, via a possibility proof: I show that there are ways that we *could* avoid the worry. In other words, in many ways this paper has extremely modest aims: it's conditional, and it doesn't purport to firmly establish many of the claims it raises. But I think you'll find it's worth your while nonetheless. Because beneath that modesty lurks a rather wild ambition: to do a bit of arcane magic, conjuring a moral ought from an action-theoretic is. It's a big project—too big for a single essay. But as with all conjurings, the first step is to catch a glimpse of the hidden realm. So that's what I'm offering: an invitation to a new way of seeing. Hopefully, by the end you'll be inclined to look at things through my eyes.

Just not, you know, in a Hannibal Lecter way.

4.1 I HOPE YOU KNOW WHAT YOU'RE DOING

Ask me why I'm putting my sneakers on, and I'll say I'm going to a march against the disaster du jour. Ask me why I'm going to the march with pico de gallo in my hair, and I'll say I had no idea I was doing that. Anscombe (1963, 9ff) says the ability to answer 'why?' questions like these—where the sense of 'why?' being deployed is “that in which the answer, if positive, gives a reason for acting”—is the mark of intentional action: so I am intentionally going to the march, but I am not intentionally doing so with pico in my ponytail. When you act intentionally, she said, you *know what you're doing*, and how, and why. That's how you're able to answer the question. (She also says you

1. That's a big project and I take it up in my (ms a).

know it in a special, direct way—‘without observation’ or inference—but that specialness won’t play a role in my arguments here.)²

There’s something appealing in Anscombe’s idea, but there are obvious counterexamples. For example, as Davidson (2001, 91–92) emphasized, I can be doing something intentionally without believing I am doing it, because I have serious doubts about my ability to pull it off. If by participating in this march I am actually somehow making a difference (despite my deep suspicion that I am howling into the void), I will not be making a difference *by accident*—it’s clearly intentional. Yet my failure to believe I’m making a difference means I certainly don’t know I’m making a difference, so my making a difference *can’t* be intentional on Anscombe’s view. Something’s gotta give.

I propose a novel way to avoid the counterexamples. I argue for it independently, not on Anscombe’s terms, but the end result is Anscombean in spirit. Anscombe makes an agent’s possession of *practical knowledge*—the special agential knowledge of what one is doing, how, and why—a necessary condition for intentional action. And that, I argue, is her fatal mistake. On my view practical knowledge is instead a *normative* condition: an agent has practical knowledge *when things go well*. In other words, intentional actions aim at instantiating practical knowledge; intentional actions are *defective* when (and to the extent that) the agent fails to know what she is doing, how, and why in this special way. So if I am ϕ -ing, but in doubt, so that I don’t know I’m ϕ -ing, things are not going well from the point of view of intentional action.

Moving from a necessary to a normative condition means I can still be ϕ -ing intentionally in a Davidsonian doubt case. When the condition is a normative one, the question of whether the agent has practical knowledge

2. See Anscombe 1963, 13–15, 49–50 on the non-observational, non-inferential nature of this knowledge; see my (ms a) for how I understand its specialness and distinctive practicality.

is a question of how we are to *evaluate* her action, instead of how we are to sort it in a binary intentional/not-intentional scheme.³

I can't provide the argument for my view here; that's the project of my (ms a). I'm not going to try to convince you; in this paper, I'm going to ask you to assume I'm right. But I can give you a sense of how the argument works, what kind of machinery I'm wheeling out, at least enough to convey the flavor of the view. First, I note that we can distinguish intentional actions from nearby phenomena—idle predictions, wishes that come true, failed attempts and lucky successes, cases of deviant causation, the things that befall us, etc.—by recognizing that intentional actions, unlike the others, are supposed to involve *control* on the part of the agent. When you set out to do something intentionally, the idea is that that thing will happen, and it will happen *because* you intended it to. Its happening will be a product of your intervention; you will have shaped the world to your will. By contrast, when we predict or wish for some outcome, we expect or hope it will happen without making any claim about our role; and when things happen to us, instead of through us, the whole point is that it was not our will that made it so.

Second, if we think about skillful actions, we see that this control is just Anscombean practical knowledge: the agent's special knowledge of what she is doing, how, and why. When Serena Williams plays tennis, she has exceptionally deep practical knowledge about what she is up to: she knows just how to hit the ball, she knows the precise moment to do so, and she knows exactly why that's the way and that's the moment to hit it. And that is to say that what happens on the court is under her control, in the sense I'm saying is relevant to intentional action, to an enormous degree.

This knowledge need not be, and often will not be, articulable or conscious. It is not necessary that Serena be able to *narrate* her skillful performance.

3. I'll say more about the relationship between aims like this and evaluation in §4.2 below.

She can be in a state of flow. What matters is that her devastating aces and her breathtaking winners and all the rest are not a matter of luck: it wasn't the wind that placed the ball there, it's the result of a precisely calibrated movement of her body which itself is the product of years of training—training that means Serena sees the conditions with an expert eye, and she responds expertly, and she knows why the response is the right one for the conditions. All together, that seeing/responding/knowing constitutes Serena's practical knowledge—which is to say, the control wielded by an agent who is doing very, very well.

Serena intends to take *this* state of affairs and make it *that* state of affairs, by doing *thus and so*. Her intention therefore involves belief: a belief that she is taking this state of affairs and making it that state of affairs by doing thus and so. When things go well, as they so often do for Serena, the intention/belief is knowledge, because it is accurate and not as a matter of luck. When things go well, what happens is what Serena's intention describes, *because* Serena's intention describes it: the intention guides the action and secures the outcome. That's what practical knowledge is—an *executed* intention. And that's what I mean when I talk about control.

By contrast, when I play tennis—always badly—I am in a state of Davidsonian doubt. I have only the smallest degree of belief that my doing thus and so will turn this dimly perceived state of affairs into the one I aim for. (I recognize that it is more likely I will serve an ace by flailing my arm toward the tossed ball than it would be if I had stayed in the locker room and refused to enter the court, but not by much.) So my action is defective, *qua* intentional action, even if I occasionally, miraculously, serve an ace—because I don't know my *thus and so* will lead from *this* to *that*.⁴ It's lucky, a fluke, when things go my way; it's not under my control in the way it is for Serena.

4. In the full argument for my view in my (ms a), I leave open the possibility of *basic* actions: actions for which no means are necessary. But I won't discuss those complications here.

What my view provides is a way to understand both what unifies my serve and Serena's—the aim of control, or practical knowledge—and the ways her serve is both different and better, *qua* intentional action, than mine.

As this discussion suggests, control comes by degrees, so practical knowledge does, too. For intentional actions to be distinguished by an aim of control—practical knowledge—is for them to be subject to a *graded* standard; they are defective when and to the extent that the agent fails to achieve the aim. I've already mentioned one way an agent can fail to have practical knowledge: she might doubt that she is doing what she intends, because she is unsure about her ability. But there are other ways. If we think of intentions as having the form *I will ϕ by means m for reason(s) r* , then in order to know the whole complex an agent has to know a bunch of other propositions: she has to know the facts that serve as her reasons; she has to know that her means are such that, if she takes them, she will ϕ ; and so on.

But the agent might have an attitude toward those subsidiary propositions that falls short of full belief, and thus holds her short of full practical knowledge in just the way Davidsonian doubt does. She might also know these propositions in more or less detail—and the extra detail makes for more control. (I know a clay court responds differently than grass. Serena knows all about *how* differently it responds. And that knowledge allows her to leverage those differences to control her shotmaking.) So the gradedness of control can manifest in different degrees of confidence in propositions involved in an intention (about one's circumstances, one's means, one's reasons, one's ability, etc.), and/or in differences in the level of detail of the agent's knowledge, and so on.

In short: intentional action is distinguished by an aim of control, and it turns out that that control is just practical knowledge in Anscombe's sense—the special knowledge agents can have of what they're doing, how, and why. When I'm acting intentionally, things go well when and to the extent that I have control, which is to say, practical knowledge. In the good case, my

intention describes what happens, and not as a matter of luck: it describes what happens because it *shapes* what happens. In a principle:

NEED TO KNOW: When you act intentionally, you aim at control, i.e., practical knowledge of what you're doing, how, and why; when and to the extent that you fail to have such control/knowledge, your action is *defective*, though it can still be intentional.

That's all I'm going to say about how I get to my view about the nature of intentional action. And again, I don't take myself to have said enough to convince you it's correct; my goal has been just to give you a sense of how it works. Now, for the rest of this paper, I'm asking you to just assume that I'm right. Because if I am, some interesting—and worrying—things follow.

4.2 PRACTICAL KNOWLEDGE CONSTITUTIVISM

NEED TO KNOW is the metaphysical hat from which I'll pluck a metaethical rabbit.

By which I mean: the principle gives us the materials for a new kind of constitutivism about practical reason; it enables us to derive an ethical *ought* from an action-theoretic *is*. Set aside for now just what the ethical content will be—it will be ethical in a wide sense in that it concerns what one is to do, but it might or might not involve, say, a concern for the welfare of others. We can't tell in advance what the *ought* of practical knowledge constitutivism will require; for now I want to focus on how we generate an *ought* at all, to show that this kind of derivation isn't *prima facie* fishy. (Look: nothing up my sleeves!)

All constitutivists are in the hat-rabbit business, arguing that some defining feature of agency generates the standards of practical reason. Following Setiya (2013, 2–4), we can schematize the constitutivist's argument. The first premise goes back to Aristotle: when a kind has a defining function, to be a good instance of the kind is to perform that function well. A good

thermometer measures temperature well. A good kidney filters blood well. This is a modest claim, and it's hard to deny while maintaining a coherent conception of a defining function.

Constitutivists say that agency, like thermometers and kidneys, has a defining function—so a good agent performs that function well. Different constitutivists disagree about what that defining function is, but they agree that there is one. Typically, constitutivists construe the function of agency teleologically: agency *aims at some end*. Velleman (2000) says the aim is self-understanding; Korsgaard (2009) says it is self-constitution. End talk allows us to cash out what it is to carry out agency's function well: it is to achieve the end *in full*.

Importantly, to play the needed role in a constitutivist argument, achievement of the end must be graded. If it is all or nothing, the aim will sort things by whether or not they are instances of agency, whereas a graded aim allows us to sort instances of agency along a spectrum of better and worse.

So the first two premises of a constitutivist's argument are the claim that agency has some defining function that's achievable by degrees, and the claim that when something has a defining function, a good instance of the kind is one that performs that function well, i.e., in full. One more premise is required: the claim that to be a good instance of agency—one that achieves its defining function in full—is to satisfy the requirements of practical reason.

Is it true that if an instance of agency is good as such, then it satisfies the requirements of practical reason? As Setiya (2013, 4; 2014, 71) notes, it's hard to see how it could be false. The requirements of practical reason have to do with the standard of excellence for actions, or agents, or instances of agency. Certainly, one might wish to deny that the *basis* of the requirements of practical reason is to be found in the nature of agency and functional kinds. But if the requirements of practical reason are not about the standard of excellence for actions/agents/instances of agency, what on earth is their subject

matter? The thing is, once you concede that practical reason has this subject matter, it's hard to see how you can deny the equivalence. A good instance of agency—which is what the first two premises purport to describe—*must* be one that satisfies the requirements of practical reason.

Let's lay out the schema:

P1. A good instance of a functional kind performs the kind's function well, i.e., (if the function is graded) *in full*.

P2. Agency has a (graded) defining function, which is x .

→ C1. So to achieve x in full is to be a good instance of agency.⁵

P3. To meet the standards for being a good instance of agency is to satisfy the requirements of practical reason.

→ C2. So to achieve x in full is to satisfy the requirements of practical reason.

A key thrust of Setiya's (2013) is that this form of argument unites constitutivists; it's logically valid; and at least premises 1 and 3 are hard to deny. If you want to avoid the ultimate conclusion, you have to overturn the second premise. So I won't argue further on the schema's behalf; I'll just point out that if I'm right about intentional action, as I've asked you to suppose, I've earned a seat in this apparently seaworthy boat.

How so? Well, what's needed to generate a constitutivist view is material for premise 2: the claim that there is some x that is the defining function of agency, plus a specification of that x that supplies an end achievable by degrees. My view of the nature of intentional action gives us just what we need—I don't talk about agency, but rather about intentional action, as the thing with a defining function, but this adjustment doesn't change

5. If this is Velleman's version of the argument, x =self-understanding. If this is Korsgaard's version, x =self-constitution. So Vellemanians think good instances of agency are ones in which the agent has full self-understanding; Korsgaardians think good instances of agency are ones in which the agent fully constitutes herself.

the shape or validity of the argument. I specify this defining function as the provision of control, which is to say, the achievement of practical knowledge. And I've already said that control, i.e., practical knowledge, is achievable by degrees.

I've asked you to suppose that I'm right about intentional action; now you see why. If I am, then my view generates a form of constitutivism about practical reason: it supplies a version of premise 2, and as Setiya shows, that's what's needed to compel the constitutivist conclusion. So my view about the nature of intentional action entails a form of constitutivism. If I'm right about intentional action, as we're supposing, then what practical reason requires is that actions fully instantiate practical knowledge.

Should we welcome this conclusion? Well, like all constitutivist views, mine gives us a way to get from an *is* to an *ought*, which is no small thing. Furthermore, my view has some appealing differences compared to the best-known alternatives.

Consider Velleman's (2000, 2009) view. He likes to invoke Anscombe, and often talks about agency's aim in term of knowledge. So one might think his view and mine are similar. But in fact, the defining function of agency on Velleman's view is the achievement of *self-understanding*. He thinks things go well for me, *qua* agent, if my action is (fully) intelligible to me in light of my beliefs, my desires, and, crucially, my character. Does my action comport, in a robust way, with my sense of who I am? If so, Velleman says, it is a good intentional action, one that satisfies the requirements of practical reason.

But as Bratman (1991) notes, if Velleman is right, then a nervous actor, who expects her stage fright to make her flub her lines, *does well* if she flubs her lines—because her doing so is highly intelligible to her. Likewise, Velleman's view suggests that a habitual procrastinator can achieve complete agential success by failing to finish writing her book, since her procrastination is in keeping with her character. And it seems a racist, corrupt sexual predator will satisfy the demands of practical reason if he just keeps on keeping on.

On Velleman's view, practical reason never demands that I be anything but what I think I already am, or do anything but what someone like me would do. But that seems very strange.

I think that Velleman, though he's right that the function of agency has to do with a kind of inquiry, is confused about the object of that inquiry. He thinks that the inquiry is inward-looking: Velleman's central thought is that whenever I act, no matter what else I'm after (a fresh mango, a tenure-track job, health), what I'm really seeking is self-understanding. No matter what my particular project is in a given instance of acting, the deeper, constant, true project is about *me*. For Velleman, the aim that defines agency is *narcissistic*.

My view is different. Even if you read Velleman's view in the way that sounds the most like mine (taking Vellemanian self-understanding talk to just be about self-knowledge), and read mine with practical knowledge's status as a kind of self-knowledge in mind, our views point to different objects of inquiry.⁶ For Velleman, the object of inquiry is the psychological explanation of one's action. For me, it's the facts in light of which one is acting.

Here's an analogy. Consider two people who decide to learn about their ancestry. One, call him D, wants to find out what his genetic makeup is like: is he predisposed to certain diseases? Does he have Neanderthal DNA? The other, call her A, wants to learn about her ancestors: were they from Canada or Cambodia? Were they bakers or biographers? There's a sense in which we can say D and A have the same project: to learn about their ancestry. But there's a difference.

The object of D's inquiry is himself; he's interested in the features of his

6. And as I said above, Velleman in fact thinks that something stronger than knowledge is required—he's talking about a rich kind of self-understanding. This is very different from practical knowledge in Anscombe's, and my, sense.

genome. Of course, in learning about himself, he will learn some things about other people—such as that someone in his ancestral line had the mutation on their X chromosome that causes red-green color blindness, or that someone in his ancestral line was a Neanderthal. But the things he learns about other people are not the things he’s aiming to learn; they’re incidental.

What about A? Well, the object of her inquiry is other people; she wants to know where her great-grandparents were born, and what kinds of lives they led. Of course, in learning about them she will learn some things about herself—such as that she’s part Canadian, or that she’s partaking in a family tradition by becoming a baker. But the things she learns about herself are not what she’s aiming to learn; they’re incidental.

The difference between D and A can seem subtle; it’s a difference of focus. But it matters: D’s project means he’s not interested in certain information (say, about which of his ancestors was a Neanderthal, and when the colorblindness gene entered the family tree). A’s project means she’s not interested in the biological facts about her genome. These differences will call for D and A to take different steps in their inquiries; they’ll rationalize different approaches. D and A have different reasons.

My view about intentional action differs from Velleman’s picture of agency in much the way that A’s project differs from D’s. D, like a Vellemanian agent, is inquiring about himself, his biology. D’s project is narcissistic. But A, like my agent, is not inquiring into herself. It’s true that the knowledge she gains will involve her, but her questions are not inward-looking. The aim of practical knowledge I articulate is, remember, an aim of control—it’s about reaching out into the world and intervening.

For my part, I also find it strange to suppose that the fundamental project of agency is about *the agent*; it just seems more plausible to me that agency’s

objective has to do with what the world is like.⁷

There is of course a vast literature on the varieties of constitutivism (and its discontents). I can't respond to all of it here. What I hope is clear is that my view offers a novel answer to the question of what the defining function of agency is versus the most famous alternatives—I say the aim is control, i.e. practical knowledge, rather than Vellemanian self-understanding or Korsgaardian self-constitution. And I think the aim I propose is more plausible. Agents may learn about themselves, or even constitute themselves, by their actions. But I don't think that's the point. The point of acting is to shape the world.

So I'm offering a new kind of constitutivism. Like any constitutivist view, it has the advantage that it shows us how normativity can arise from non-normative stuff. Furthermore, my view is more intuitively plausible than its most famous cousins. So should we welcome the news that if I'm right about intentional action, then what practical reason requires is just that actions fully instantiate practical knowledge? Unfortunately, there's a reason to think the answer is no.

4.3 WHO'S AFRAID OF PRACTICAL KNOWLEDGE CONSTITUTIVISM?

Why might the truth of my brand of constitutivism be unwelcome? The answer comes in two parts. The first is easy to see, though it will be hard to see off. The second is harder to see, and I think can't be seen off at all.

So, the first part: it's not at all obvious whether or how a requirement of practical knowledge entails all the other requirements we think good actions

7. While I don't have space to discuss the point in detail here, it's worth noting that this point applies equally to Korsgaard's view, on which actions are supposed to create agents. One might have thought it was the other way around. I suspect a Korsgaardian variant of narcissism may underlie the famous shmagency objection from Enoch (2006), but I don't have space to explore the point here.

must satisfy. We're supposed to be deriving all the standards of practical reason from the nature of intentional action—but how would the derivations go for particular principles? How do we get, say, a requirement of means-end coherence out of the requirement of practical knowledge?

I think this one is doable; just think about how the difference between Serena's and my knowledge about our means in playing tennis figured in the discussion above and you'll have a sense of how it might go. (Agents only have full practical knowledge when their means *secure* their ends. If the means you intend will not in fact lead to your end, then if your end comes to pass it won't be *because* you intended it—it will be lucky, instead of a product of your control. Serena's deep knowledge of her means—what they are, how they work, how they relate to her circumstances—is why she is such an exceptional agent, with such extreme control.)

But what I want to focus on is something that looks much harder to get out of the materials provided by NEED TO KNOW. How could we get classically moral requirements like concern for the needs and interests of others from the requirement of practical knowledge?

Now, you might think that we just won't. What practical knowledge constitutivism tells us is what it takes for actions to be good as such, but we shouldn't expect to get all our normativity that way. An action that's good as such may or may not be morally good, we might think. But this is where the second part of my answer comes in.

Constitutivist views are powerful: they summon ought from is. But like all powerful magic, they exact a price. Constitutivist views don't just offer to provide normativity; they insist on it, and they won't tolerate your shopping elsewhere. This is the force of the identification in premise 3 above. (Once you've said that what it is to be a good thermometer is to measure temperature well, you can say that thermometers ought to measure temperature well. But you can't say thermometers ought be, say, amusing. Amusingness or lack thereof just doesn't bear on what thermometers should be like, if measuring temperature is the defining function of thermometers.) As Setiya

(2014) puts it in arguing for this point at length, constitutivist views are *imperialist*.⁸

Let's walk through this a little more slowly. Premise 3 identifies being good qua exercise of agency with meeting the standards of practical reason, because the standards of practical reason just are the standards of agency. Given the rest of the argument, if there is a graded constitutive aim of agency, then that's where we can *and must* look for the basis of the requirements of practical reason. But this means that any form of constitutivism threatens to limit the existence and/or generality of reasons.

This is because one way to fall short according to the standards of practical reason is to fail to respond to a reason of which one is aware. But if some form of constitutivism is true, and one can fail to respond to some fact of which one is aware *without* failing to satisfy the constitutive aim of agency, then that fact is not a reason.

This adds new weight to the worry raised above about whether and how we might get a requirement of concern for the needs and interests of others out of a requirement of practical knowledge. Suppose I see someone suffering terribly, and I see that I can alleviate that suffering at no cost to myself. If I can intentionally refuse to alleviate that suffering, being utterly unmoved by it, and do so with full practical knowledge, then the fact that that suffering exists and I can remedy it at no cost to myself is not a reason to remedy it.

This is a consequence of the way my view of action, if it's right as we're supposing, compels a constitutivist view, combined with the role premise 3 plays in the constitutivist argument schema. Again:

Premise 3 identifies being good qua action with satisfying the requirements of practical reason. If you are satisfying the requirements of practical rea-

8. I follow his argument in what I say here.

son, you are not failing to respond to a reason of which you are aware. Therefore the reasons you have are all and only those such that, if you are aware of them, you cannot fail to respond to them while fully achieving the constitutive aim of action.

It can be hard to get a handle on exactly what the problem is here, so I want to distinguish the worry I'm raising from one that might seem similar. I'm saying that if practical knowledge constitutivism is true, then the only reasons we have are those that come from the aim of practical knowledge. But here's what this *doesn't* mean: it doesn't mean that if I have a choice between accepting five dollars and shooting at a bullseye with a 75% chance of hitting it, where hitting the bullseye comes with a prize of five hundred billion dollars and an end to all human suffering, I have to take the five bucks. The truth of practical knowledge constitutivism doesn't mean that what you have most reason to do is what you'll have the most practical knowledge in doing; it's not an assimilation of all possible reasons to the single reason *I'll know I'm doing this*.⁹ That way of thinking is a category mistake: the achievement of practical knowledge isn't itself a reason, it's a source of and formal constraint on reasons.¹⁰

So if I'm right about action, we'd better hope that, somehow, you can't really have control, you can't have full practical knowledge, without being moved by the needs and interests of others. (Yes, I'm looking at you, Mr. Lecter.)

The question, then, is obvious. How could a requirement of practical knowledge generate a requirement of responsiveness to moral considerations? To put a finer point on it: doesn't Hannibal Lecter know quite well what he's doing? Doesn't he have quite a lot of control? It seems natural to think that when he manipulates everyone around him to enable his escape and resume

9. Thanks to an anonymous reviewer for pressing me on this point.

10. For discussion of another kind of comparative question, see §4.7.3.

his brain-eating ways, he knows *exactly* what he's doing. So the question is, what resources does practical knowledge constitutivism have if we want to criticize Hannibal Lecter, to say he's not actually the ideal agent? How can we tell him he's going wrong in not responding to moral reasons?

Call the worry that practical knowledge constitutivism doesn't have robust enough resources to allow us to criticize a monstrous cannibal serial killer the *thin gruel* problem. In the rest of this paper, I'm going to suggest a couple of ingredients we might add to our constitutivist stew to make it into a more satisfying dish.

4.4 ONE WAY OUT: OOMPHY CONCEPTS

Our first possibility borrows an idea from Murdoch. She says if you are fully competent with the concepts you use in describing your circumstances, you'll be responsive to the needs and interests of others in the ways that are morally appropriate. To know a fact is to be moved by it, in just the way and to just the degree morality calls for.¹¹ As she puts it (1970, 64), "true vision occasions right conduct." So if you really know the facts, you'll do the right thing—because really knowing requires full competence with the relevant concepts, and full competence with the relevant concepts means being responsive to features like the needs and interests of others. Since really knowing the facts is just what practical knowledge constitutivism requires, then if Murdoch is right, we have a solution to thin gruel. Full competence with the concept of *need* or the concept of *interest*, or the concepts that go into describing needs and interests, will entail responsiveness to those features of your circumstance. And so actions that achieve the aim of practical knowledge will be actions that are morally right, in the sense that they manifest appropriate concern for, or responsiveness to, the needs and interests

11. I owe this simplified statement of Murdoch's view to Setiya's (2013, 7) reconstruction; what I say about her view will almost entirely agree with, and is deeply indebted to, his reading.

of others.

The trouble is that it's easy to think of people who, say, know that millions of people are starving in Yemen but still think it's okay to spend half a billion dollars on a possibly fake and certainly damaged painting by Leonardo da Vinci. So why would we think Murdoch is right?

Murdoch anticipates this objection. She replies that such a person doesn't really, fully, know. You may think they know, but that's because you're confused about how concept possession works. Sure, this person is competent enough with the relevant concepts to pass ordinary public tests—the unmoved art collector, we'll assume, could give a definition of 'starving' and 'painting' and carry on a conversation about the purchase and the famine that isn't obviously incoherent. But Murdoch says that this isn't all there is to concept possession; this kind of ordinary public competence is a lowest common denominator, not an outer limit of possibility. That we can do better than mere public competence is the point of her famous discussion of the case of M & D.

As the story begins (Murdoch 1970, 16–17), M thinks of her daughter-in-law, D, as “pert and familiar, insufficiently ceremonious, brusque, sometimes positively rude, always tiresomely juvenile.” But M engages in some reflection, attending more carefully to D, and comes to see her differently. “D is discovered to be not vulgar but refreshingly simple, not undignified by spontaneous, not noisy but gay, not tiresomely juvenile but delightfully youthful, and so on.”

The point is that, in reflecting, M improves her grasp of the concepts she applies. She begins with ordinary public competence, but her attentive effort enables progress beyond that. She comes to see that the concepts in her initial characterization of D were the wrong ones, and to determine which concepts are actually appropriate, by gaining a new grasp of what both sets of concepts involve.

To be clear: Murdoch is *stipulating*, not arguing, that M is getting things right after reflection that she wasn't before. The point of the example is

that this story is *possible*; people can start out with ordinary competence and then improve beyond that.

If Murdoch is right about M & D, then apparent counterexamples to the claim that really knowing a fact means being moved by it as morality requires need not be counterexamples at all. It can be true that a person who is unmoved by a circumstance in which she can save a life at no cost to herself is not seeing the situation clearly. Such a person, we can say, is not fully competent with the relevant concepts, and so does not fully know what her circumstances are.

So far, this is just a possibility proof.

How can it help with practical knowledge constitutivism? Recall that the reason NEED TO KNOW can serve as the basis for a constitutivist argument is that it provides a standard that can be met by degrees: an agent, in acting intentionally, can have more or less control, which is to say more or less practical knowledge.

As I said in the initial articulation of my view, we can have more or less practical knowledge in virtue of differences in confidence, or in virtue of knowing more or fewer of the propositions relevant to what we're doing, how, and why. But as the discussion of Murdoch brings out, there's another way our practical knowledge is graded: we can have more or less grasp of the concepts that figure in those propositions. Because concept possession comes by degrees, so can our knowledge of particular propositions. Young children learn that gravity makes baseballs fall back to earth no matter how hard you fling them into the heavens. They know this fact, in the sense that they are competent in at least a minimal version of the public sense. But AP Physics students in high school know the same fact better, and professional physicists know it better still.¹²

12. See my (ms b) for more about how differences in conceptual grasp affect what one can know and thus what one can do.

The case of M & D shows that full possession of a concept is not merely a matter of passing the tests of public performance, and the phenomenon is eminently recognizable. New parents often say they only now really *get* what love is. A soldier discovers what war is when she experiences combat for the first time. It's not that the parents and the soldier had not achieved public competence with the relevant concepts before the arrival of the child or the trip to the front lines. It's that there's more to grasping love and war than what mastery of public use requires.

So the fact that we might be inclined to think someone can know a proposition without being moved by it is just a reflection of the shallowness of the requirements of public competence. Once you see that we need not equate an understanding that deploys ordinary public competence with one that manifests the deeper kind of knowledge M achieves, it is not, after all, outrageous or impossible to think true vision occasions right conduct. (The parent and the soldier will attest that their new knowledge is loaded with action-guiding oomph.) But it's of course a big step from 'it's not impossible' to 'it's true.' How are we supposed to make that step?

Murdoch's view is that the progress M makes is an instance of a generally possible deepening and enriching process, which could be perfected. What would a move toward perfection be like?

Competence with a concept is twofold, so perfect competence is, too. For full competence, one must not only be disposed to recognize or categorize the *F*s as *F*s—applying a concept is not merely a matter of naming. Concept application also, essentially, involves a disposition to respond to *F*-ness in certain ways. I'm not competent with *hunger* unless I both recognize it as hunger *and have a disposition to get something to eat* when I'm experiencing it. The kinds of responses that follow from the identification of hunger include both inferences (I'll conclude that I need some food) and motivations (I'll be moved to get some food). In short, what's needed is compliance with both what we might call *input rules*, which determine which things count as *F*s, and *output rules*, which determine the appropriate responses to *F*-

ness.

On this way of thinking, *perfect* mastery of a concept, *full* possession, will be like perfect skill in other domains: being disposed to do just what you should do, just when you should do it, across all circumstances—calling all and only the *F*s ‘Fs,’ and doing everything called for when in the presence of *F*-ness. If that’s right, then we’ve made significant progress in dispatching *thin gruel*: potentially robust motivational, practical content can come with the requirement of practical knowledge, because achieving full practical knowledge will require perfect concept possession, and perfect concept possession requires certain responses.

So we’ve made one more step: we’ve shown that true vision occasions... conduct. But to fully dispatch *thin gruel*, we need yet another step, to make sure that the responses required will include concern for the needs and interests of others.

Murdoch handles this by saying that we will respond to the needs and interests of others in the ways that are morally appropriate *because those are the reasons*, and nothing could be a reason if it didn’t mandate moral conduct. But this is where I part ways with Murdoch.

I need to, for one thing: if I went Murdoch’s way, my view would be circular. (I explain reasons in terms of the aim of agency, so I can’t appeal to the notion of reasons here; doing so would be explaining the aim of agency in terms of reasons.) So I want to agree with Murdoch that concepts like *need* or *killing* come with output rules, without taking on board her explanation for why those rules prescribe concern for the needs and interests of others.

That means I still need an explanation for why full competence with *killing* will involve avoiding it, full competence with *need* will involve ameliorating it, etc. Could one have a concept of *pain* that is non-aversive, or non-aversive when it belongs to others and aversive only when it is one’s own?

I can’t say everything I’d like to about this here. (Which is one reason I

promised only a possibility proof in this paper.) But we can make enough progress to see that the possibility proof goes through.

Start here: A non-aversive concept of *pain*, or a concept of *killing* that is non-aversive when the target is someone else, may or may not be possible. But the concepts of *pain* and *killing* that are actually in use are ones where pain is aversive, and killing is bad. So even if there were no further anchor (such as the one Murdoch finds in her Platonic metaphysics), as long as we're accountable to the actual public concepts, we'll be making an error if we are aware our circumstances involve killing or pain and we are not moved to avoid them. Our public concepts of pain and killing have these roles, and we are accountable to the public concepts. Why? I take it that that this publicity of one's concepts is the price of entry for communication and coordination—so there is an important constraint placed on our concepts by the social world.

So if we imagine some person who is not using the public concepts, well, that person may fail to know things relevant to her action—say, that it is a killing, if *killing* is not among her concepts. But then we will just evaluate the action according to our thinking about agents acting in ignorance—there is no special problem here.

If we press on, wondering, couldn't there be a genuine concept, *killing**, which has the same input rules as *killing*, but output rules such that *killing** is non-aversive? Then even if *killing* is also a genuine concept, it looks like there would be no verdict provided by the knowledge that one's action was a killing/killing*.

I think this is unlikely, partly for reasons I'll mention in §4.6. But also, nothing I've said here entails that the normative weight of applicable concepts can't be in tension; we might think that a case like this would be a somewhat unusual version of that possibility.

To summarize: as long as one is fully competent with the *actual* concepts of pain, killing, etc., one will be moved, as Murdoch says. If one was using different concepts, one might fail to know important facts, but we can deal

with a case like that in the usual way. And if we wonder whether there could be genuine concepts that had the same input rules but wildly different output rules, well, it's not obvious that that's possible, or exactly how bad it would be if true. In any case, we've arrived at a point where the only way *thin gruel* could retain any bite at all is if there are coherent, genuine concepts like, say, a non-aversive *pain**, such that an agent who is using such concepts is not in conflict with her social environment, the physical/scientific facts, etc.

Given all that, I take it we have one promising potential way out of our worry about moral nihilism.

4.5 ANOTHER WAY OUT: KNOWING ALL THE WAY

Maybe your eyes glaze over at concept talk. Maybe you're really worried about *killing**. If so, I'm happy to report that there is another possible avenue of escape from *thin gruel*.

Remember, practical knowledge constitutivism says an action is good to the extent that the agent has practical knowledge, and defective to the extent she does not. So if you could know more fully what you're doing (and how, and why), then your action is defective. And of course, one way you can fail to know fully what you're doing is to fail to know, fully, some fact about your circumstances, your means, your reasons, or the outcome, as those things figure in your intention.

What would it mean to know such a fact, but less than fully? I'll make one way of understanding the thought more concrete in §4.6 below, but for now I want to give you a sense of what I have in mind in another way. Recall Jackson's (1982) discussion of Mary the scientist. She knows all the facts about color, but has never had the experience of seeing red. Jackson's claim is that she learns something when she sees red for the first time: finding out what it's like to see red enhances her knowledge.

We can understand her learning as either the acquisition of some new fact, or as a deepening of her knowledge of a fact she already knew; I'll say

more about what follows depending on which we choose. What interests me now is that, either way, we see that knowing *what it's like* is very often an important way of knowing *more fully*.

Now set Mary's case aside for a minute, and just think about the difficulty you have in deliberating when you contemplate taking action that will involve your having a really new experience. You can think here of the kind of profound life changes discussed in the transformative experience literature, but I'm going to use a more mundane case.

You're considering taking up hot yoga. The classes are expensive, but you've got some physical issues that make other kinds of exercise unsuitable, and you've been told hot yoga is your best option. There's a discount on the classes if you pay for ten up front; at that price, you could swing it. But you normally dislike heat, and you usually prefer fast-paced activities like basketball. So you're not sure what it will be like for you to take the classes, and if it's bad in ways that make you decide to quit, it will be a big waste of money that you can't afford.

What you're wondering is if the experience of hot yoga will be bad in the ways you fear, and, if so, if it will be bad enough to make buying the ten classes a waste of money. You'd know more about what you were doing in buying the classes, and in a way that is significant for whether your action is defective by the lights of NEED TO KNOW, if you knew what it would be like to take the classes.

The point is that often one of the challenges we face in deliberation is that we don't know ahead of time something that's important for our intentions: what it will be like to take the action in question. We can't reason as well, in forming an intention, if we don't know not only that (for example) it will be over a hundred degrees in the yoga studio, but also *what it will be like* to do yoga in a studio where it's that hot.

Intentions formed in deliberative circumstances like these will be defective, ineligible for full practical knowledge, because they'll either take a stand on questions the agent doesn't have evidence for, or they'll remain agnostic,

and thus fail to provide guidance, about significant points.

But notice: we can say the same thing about our knowledge of what-it's-likeness regarding features of our action when the person who'll have the experience is someone else. Very often, another person's phenomenology will figure in our intentions. If I have to break some bad news to my friend, I'll wonder if it will be less painful for her if I tell her quickly or build up, if I do it over drinks in a bar or at home over tea, etc. (And if she were my frenemy, so that what I wanted was the option that would hurt the most, I'd still be better able to deliberate if I knew what each option would be like for her.)

So if your intention involves facts about what your action will be like for someone, you can't have full practical knowledge unless you know, fully, what it will like for that person—whether it's you or someone else.¹³

Now remember Mary. Even if we include facts about what it's like to see red in the list of things Mary supposedly knew before she saw red, we'll want to say that she learned something when she had the experience herself. Having an experience is the obvious, and paradigmatic, way of having knowledge of what that experience is like. Similarly, if I've had an experience before, I'll be more confident of my ability to reason about what it's like.

But notice two things about that. First, having the experience in the past normally helps me to properly deliberate and plan about it now. Second, the reason it's so useful to have had the experience is that when you're having the experience, in the moment, its normative contours are inescapable. When you're in pain, the aversiveness of pain is obvious. And having experienced a particular kind of pain enables responsiveness to these built-in normative properties in all their specificity, including the particular ways the make-it-

13. This should actually be a claim not just about people, but about any entity for which there can be something it is like. (I certainly spend an inordinate amount of time configuring my plans based on what things are like for my cat.) But I'll talk about people here for simplicity.

stopness is mitigated by distractions, or balanced by the phenomenology of benefits the experience also promises, and so on.

This is not to say that we must have had an experience, or be having it at this very moment, in order to know, fully, what it's like. Memory is not necessarily inadequate, and we may be able to imagine our way into knowledge of what some new experience will be like, especially if we've had similar experiences before. But what is required is that the proposition *that some experience will be like this* has the same motivational contours whether you're having the experience, remembering it, or imagining it. If the motivational upshots that are inescapable from inside the experience drop out, you've lost your grip on an essential feature of the phenomenology, so you don't fully know what it's like.

Let's take stock. Very often facts about what our actions will be like—either for ourselves or for others—figure in our intentions. And I've argued that fully knowing what something is like requires having a motivational profile that matches the one that is automatic when one is in the throes of the experience. So if pain is part of the *what it's like* content, your motivational profile must reflect that, including the built-in aversive qualities, and in the kind of detail that captures the specifics of the pain in question. And this will be true whether the person who will be having the experience is you or someone else: if facts about what it will be like for someone else figure in your intention, then your intention will be defective, because it won't be eligible for full practical knowledge when executed, if your action isn't shaped by the motivational oomph that is manifest from inside that experience. So if I intend to do something that will cause you pain, and either that pain or what it will be like for you in general figures in my intention, then my intention, and action, will be defective if they are not *shaped* by what it's like for you just as they would be if I plugged myself into an experience machine and felt the pain as you will.

In a nutshell: if what my action will be like for someone—anyone—figures in my intention, then my action will be defective if it is not shaped in ac-

cordance with the motivational oomph that is inescapable from inside that phenomenology. You don't fully know what you're doing in such a case if you aren't appropriately motivated. And if that phenomenology involves (e.g.) pain, it doesn't matter if it's your pain or my pain, the motivational profile is—must be—aversive, because what fixes the appropriate motivational profile is what it's like *from the inside*.

How does this help with *thin gruel*? It shows that at least in the cases where what my action will be like for someone figures in my intention, my action can only be non-defective if it conforms to the motivational profile mandated by the phenomenology. But that means that it can only be non-defective if it comports with the kinds of concerns that utilitarians (at least) take to be of interest for morality. Pleasures call us to seek them (*ceteris paribus*), and pains call us to avoid them (*ceteris paribus*), no matter whose they are. My actions are defective, because I don't fully know what I'm doing, if they don't reflect that, at least as long as facts about what it's like for people figure in my intention.

Of course, we might want more.

Most urgently, the limitation on which intentions are subject to the phenomenological constraint looks like a problem for the generality of this approach as a solution to *thin gruel*. If all I have to do to avoid the requirement of concern for the needs and interests of others is to avoid having facts about what my actions will be like for them figure in my intention, then it looks like, say, a malignant narcissist, who just never thinks about other people in that way, isn't subject to the moral demands we were hoping to secure. And couldn't anyone just gerrymander her intention to avoid including the facts that come with motivational requirements of concern for others?

One thing to say is that it's actually very hard to take interesting, substantial actions without somehow invoking other people and something about how you'll affect them. You operate within social structures; you have expectations about what will provoke a response and what will not, where sometimes you need the response and sometimes you need there to be none;

you engage in communication, which crucially depends on the participation of your interlocutors. So even if you don't explicitly think, e.g., 'Lisa will help me and Yafeng won't interfere, because what I do will be like this for Lisa and like that for Yafeng' your calculation of risks and expected values and your expectations about what means are available and how they work and a host of other considerations will bring in other people. And what I want to say is that that is plausibly enough to open the door.

Next, the cases where it seems most plausible that I could leave out facts about what my action will be like for others are ones where the action in question is something like *blinking 3 seconds from now*. But that kind of case also seems like precisely the kind in which we might be inclined to say that there *aren't* morally relevant effects on other people.

Finally, remember that when I first mentioned Mary's case, I said we can understand it in either of two ways: either she comes to know a new fact (or facts) when she learns what it's like (from the inside) to see red, or she comes to know the same fact(s) *better*. If we choose the second option, we might have a way of extending the reach of the phenomenological argument about action. Here's the idea. If we think that Mary, when she sees red for the first time, doesn't just come to know better what red is like, but also comes to know better *what red is*, we might be able to make a similar move here. Maybe knowing what it's like is important for full knowledge of lots of things—we might use this move to recover the requirement of concern if an intention includes facts about killing, for example. We might even extend the thought to say that knowing what it's like is required for full knowledge of any state of affairs.¹⁴ While I can't prove this here, it doesn't strike me as outlandish.

Secondarily, as I said, this picture most obviously captures the kinds of needs and interests that utilitarians emphasize: pleasures and pains and

14. If this sounds like a way of expressing the Murdochian ideas of §4.4 in a material rather than a formal mode, I'm okay with that: see §4.6 below for more.

their cousins. What about other things that we might think are relevant if we want to capture all that counts as concern for the needs and interests of others? What about, say, rights? Of course, some people will think we don't need to capture those. Furthermore, as I said at the beginning, I am not here attempting to show that we can capture everything we might have in mind when we talk about moral requirements. It may take significant further argument to get a requirement of concern for rights from the phenomenological approach, if it can be done at all. But for now, we'll have made very substantial progress in solving the thin gruel problem if the phenomenological argument has secured a utilitarian moral framework. If that turns out to be all we can get, we'll have to decide if that fact counts in favor of a basically hedonistic view, or if it's a point against the truth of practical knowledge constitutivism, or if it means that practical knowledge constitutivism is a good news/bad news scenario. No matter what, though, *thin gruel* is substantially defanged if we secure a requirement of concern for pleasures and pains, wherever they're found.

What's more, I think that we need not suppose that the phenomenological approach only yields concern for pleasures and pains. Phenomenology has more diverse motivational upshots than that. Edges in our visual field compel our attention. The literature on affordances reminds us that a door handle says *grab me*. A full exploration of the normative content of phenomenology is the project of another paper, but it's clear there is much more to talk about than just pleasure and pain, and that this may well be a source of richer demands.¹⁵

The phenomenological approach says that fully knowing what you're doing requires having a motivational profile that matches the one you'd have if you were experiencing the effects of your action, all of them, from the inside, and that that profile will be one that is sensitive to the needs and interests of others in at least the sense that utilitarians care about. You might think

15. I thank Mark Schroeder for this point.

of this as a kind of extreme Humean sympathy. As with the Murdochian approach above, questions remain about exactly what the limits of this solution are. But given all I've said, I take it we now have a second potential escape from the *thin gruel* problem.

4.6 BETTER TOGETHER

I've offered two possible ways to avoid the conclusion that a world in which practical knowledge constitutivism is true is a world in which moral nihilism is, too: the Murdochian approach and the phenomenological approach. Either one, I've argued, could give us what we need.

I've also noted that with each, we might wonder about certain limits. With the Murdochian approach, it wasn't clear if one could have a genuine concept, *killing**, such that instances of *killing** were non-aversive. With the phenomenological approach, it wasn't obvious how many intentions would be subject to the requirement, because we might wonder if only intentions that explicitly include facts about what one's action will be like for others fall under its scope, and it also wasn't clear if that approach would ever be able to capture something like rights.

What I want to note here is, first, that the Murdochian approach and the phenomenological approach can both be right; and, second, that if they are, each can help address the possible limits of the other.

If both are right, phenomenology could provide the kind of anchor we want to constrain what concepts can be genuine on the Murdochian picture. There couldn't be, for instance, a genuine non-aversive concept of pain, because aversive phenomenology is essential to pain; if you omit the aversive phenomenology, you might say, you're talking about something other than pain, changing the subject. The general thought would be that there are some central, morally relevant concepts that are *locked in* by the phenomenology. This would provide a constraint to replace the one lost when I abandoned Murdoch's explanation for the moral content of output rules.

If both are right, the Murdochian picture could give structure to the phenomenological view, and fill in some of the areas that are hard to secure via phenomenology alone. This seems likely to help us extend beyond pleasures and pains, and perhaps beyond facts that are explicitly about what things are like. It might even enable us to capture things like rights.

Finally, I think that the combined picture puts meat on the bones of one of the most compelling things Murdoch says (1959/1997, 215): “Love is the extremely difficult realization that something other than oneself is real.” Combining phenomenology with the claim about concept possession suggests that what it means to consistently apply a concept like *pain* is for the pains of others to be as real to you as your own. In full generality, this requirement is nothing less than a requirement that the lives of others, their weal and woe, matter to you as if they were your own. And that, we might think, is love.

Because I find both the Murdochian approach and the phenomenological approach plausible, and because the combination is appealing in these additional ways, my own instinct is to say that the best solution to *thin gruel* will be one that involves the combined view. Those who found one more appealing than the other need not take on the combination, but for my part it looks to be both the most promising and the most intuitive option.

4.7 LOTS OF LINGERING QUESTIONS

I’ve been arguing that even if practical knowledge constitutivism is true, so that all there is to say about acting well derives from the aim of practical knowledge, it is at least *possible* that we can salvage a requirement of concern for the needs and interests of others. I’ve sketched a few ways we might do so. But even though I’ve claimed only to be demonstrating this possibility, not showing that it holds, you probably still have a lot of questions. I’m going to try to gesture at some answers in the space I have left.

4.7.1 DOWN WITH IMPERIALISM

Maybe you got off the boat back in §4.2. Maybe you thought, okay, NEED TO KNOW tells us what it is for actions to be good for their kind, but I don't buy this claim that that's the *only* standard. Why can't I just say that even though Hannibal Lecter is doing really well qua agent, he's doing badly by a separate, moral standard?

There are two things you might have in mind with that question. On one, you're conceding that whether Lecter is practically rational or irrational is fixed by whether he meets the standard of NEED TO KNOW—you accept the identification in premise 3 of the constitutivist argument—you're just saying that even if he can rationally (by the standard of NEED TO KNOW) fail to be concerned with the needs and interests of others, nonetheless there are reasons for him not to be that way.

But this version of the objection is incoherent. Accepting premise 3—the identification of being a good instance of agency with satisfying the demands of practical reason—is accepting that reasons come from the standard of action. That was the point of the discussion of imperialism and amusing thermometers in §4.3.

On the other hand, maybe what you mean is that whether or not Lecter is rational is beside the point; the question is whether he's *vicious*.¹⁶ This version of the question isn't addressed by the discussion of the imperialism of constitutivist views above. But I think there are things to say.

First, notice how clearly authoritative the kind-goodness standard, identified with practical rationality, is. We understand why a kind ought to fulfill its constitutive aim, why that ought has bite. And if we're conceding the existence of a standard like that, and positing a separate moral standard, there's a real puzzle about how morality could be authoritative too. Because,

16. This version of the objection is in the spirit of Williams (1979).

of course, we can come up with lots of other standards: etiquette, the rules of games, arbitrary conventions, æsthetics. If morality is just one among many of those, what makes it special? We want morality not only to condemn, but authoritatively, perhaps even overridingly. But how could that be?

One way to see the problem is to realize that if we're conceding that Lecter is practically rational, we can't really say he has *reasons* not to be a brutal killer. But even if we want to try to wriggle out of that problem, I think there's another one. Even if there's some (free-floating) sense in which Lecter shouldn't go around eating people's brains, that's not all we want. If morality is going to matter in the ways it's supposed to, we also don't want to say that there's some important sense in which he should.

In short, even the coherent version of this worry doesn't point to a way that something that really has the features we expect to find in *morality* could be free-floating.

4.7.2 EXTRA! EXTRA!

Maybe you're noticing that there is an awful lot of machinery here. Couldn't we streamline things a bit? I mean, why do we need practical knowledge constitutivism if we have the Murdochian and/or phenomenological stuff? Why not go with a simpler view, maybe a neo-Humean kind of thing on which what I have reason to do is what I would be motivated to do if I were fully informed?

There are a couple of things to say here. The first is to recall that my argument is that if I'm right about what intentional action is like, we're *stuck with* practical knowledge constitutivism. So then the question is where that leaves us—which is where the Murdochian and phenomenological material comes in.

And sure, it's true that the moves I make toward converting the aim of practical knowledge into a requirement of concern for the needs and interests of others end up leaning on what it takes to really know things, and so run into territory that's also friendly to the neo-Humean option. But remember:

one key benefit of the constitutivist approach is that gives us an answer to the question of *why it matters* what we'd be motivated to do if we were fully informed; as I've been saying, it solves the problem of the authority of the moral requirements. By contrast, the bare neo-Humean view is vulnerable to a shmagency-style objection: if being fully informed means I have to be good, ignorance is bliss! (At least enough ignorance to let me off the hook.) So it's not clear to me that the sparser neo-Humean view is more appealing.

4.7.3 NAIVE SAINTS

Here's another puzzle. My view says the more control an agent has, the more fully she knows what she's doing, how, and why, the better her action is. If that's so, we might wonder if savvy evildoers will do better by the standard of my view than naive saints. We've worried about whether we could criticize Hannibal Lecter for eating people's brains; the worry I'm raising now is about the *comparative* question: is Hannibal Lecter doing *better* by the lights of practical knowledge constitutivism than people we want to say are morally better even if less masterminded? Even if one concedes that Lecter doesn't have full practical knowledge, and so is criticizable, will his degree of deficiency qua agent line up with his degree of moral deficiency?¹⁷

There are many versions of this worry. We can compare a sadistic torturer to a clueless nice guy, where the torturer is (to use our Murdochian framework) really good at the input rules, and completely wrong about the output rules, and the nice guy is really bad at recognizing the input conditions, but very good at getting the output right on those occasions when he manages the input.

Or we can pit the torturer against an inarticulate saint, who has inchoate grasp of input but somehow manages to do well at output. What about an inept Samaritan, who recognizes need and aims to meet it, but is incompe-

17. I'm grateful to Tristram McPherson for pressing me on this point.

tent? Is she doing better than a bandit who sees the same situation and recognizes not the need but the opportunity to steal a wallet and get away with it, and does so?

Rather than taking the cases piecemeal, how could we answer this challenge all at once? In each case, what we want is for getting the output right to *take priority* in our evaluation of someone's degree of knowledge.

I think that's right. One relevant thought is that it seems like the *point* of input conditions is to direct us to output conditions—the point of categorization is to shape our responses. Consider a case of non-moralized, theoretical knowledge.

Tom and Tara are at the Large Hadron Collider. Tom is a layman who thought he was visiting a demolition derby; Tara is a graduate student in particle physics. They both come to believe the proposition *that was a Higgs boson*. Tom believes it because the Director General, Dr. Fabiola Giannotti, tells him “that was a Higgs boson.” Tara, out of earshot, believes it because of the readouts she sees on the instruments. Given the testimony on which Tom's belief is based, he knows that was a Higgs boson. But has no idea what follows from that. Tara, on the other hand, doubts her belief; it would be a big deal to have witnessed a Higgs boson. But she knows *exactly* what it would mean for it to have been a Higgs boson.

I think Tara is doing better than Tom in this scenario. Her knowledge about what the world is like has increased, even if her doubt means there is still room for improvement; the way her new attitude is integrated with the rest of her knowledge matters. Tom, on the other hand, might as well have been told “That was a blorp.” The only facts he gains access to are trivial—things like *Dr. Giannotti said ‘that is a Higgs boson.’* He can't tell if she was alerting him to the presence of a rare rodent or an expensive piece of equipment or a person of a particular rank in the organization. It's not clear to me that ‘knowledge’ of a fact involving a concept about whose output conditions one is totally ignorant can really be knowledge at all.

This case is a recipe. Even though a full demonstration that output competence is prior to input competence is work for another day, I take it there's reason to think that output should be prior to input when it comes to assessing knowledge, including the knowledge relevant to action.

4.8 CONCLUSION

I've shown that my view of the nature of intentional action entails a new kind of constitutivism about practical reason. I've also noted that we might not welcome this news, because if actions are better or worse just depending on how much practical knowledge the agent has, then either there are no moral requirements, or the moral requirements are derived from the requirement of practical knowledge—something that is not obviously possible. So the goal of this paper has been to show that the derivation *is* possible: there is a way to get from a requirement of practical knowledge to a requirement of concern for the needs and interests of others. In fact, I've argued, there are at least two potential paths to that result, and they can be taken separately or together. And there may well be other possibilities; one benefit of offering a possibility proof is that I'm not ruling out other good answers to the question of how the derivation might go.

Agential perfection, I have argued, binds success in knowing and success in doing both to each other and to a demand that we be alive to the needs of others—Hannibal Lecter be damned. But what does that ideal of perfection have to do with mere mortals like us? Murdoch calls our attention to the nature of our concepts because their possible perfection is connected to our imperfect state by a chain of potential progress. If, like M, we make the effort to attend to our world and our worldmates, we can change our way of seeing. And because what we see shapes what we do, to change our way of seeing is to change our world. It is not easy: it is an achievement to act well, and true vision may be rare. But valuable things often are. Our task is to seek them.

REFERENCES

- Anscombe, G.E.M. (1963) *Intention*. Cambridge, MA: Harvard University Press.
- Davidson, D. (2001) *Essays on Actions and Events*. Oxford: Oxford University Press.
- Jackson, F. (1982) "Epiphenomenal qualia." *Philosophical Quarterly* 32: 127–136.
- Jaques, A.E. (ms a) "Where there's a will, there's a way of knowing."
———. (ms b) "Shackling."
- Korsgaard, C.M. (2009) *Self-Constitution: Agency, Identity, and Integrity*. Oxford: Oxford University Press.
- Murdoch, I. (1959/1997) "The sublime and the good." Reprinted in *Existentialists and Mystics: Writings on Philosophy and Literature*. London: Chatto & Windus.
———. (1970) *The Sovereignty of Good*. New York: Routledge.
- Setiya, K. (2013) "Murdoch on the sovereignty of good." *Philosophers' Imprint* 13(9): 1–21.
———. (2016) *Practical Knowledge: Selected Essays*. Oxford: Oxford University Press.
- Thompson, M. (2011) "Anscombe's *Intention* and practical knowledge." In *Essays on Anscombe's Intention*, ed. Ford, A., Hornsby, J., & Stoutland, F. Cambridge, MA: Harvard University Press.
- Velleman, J.D. (2000) *The Possibility of Practical Reason*. Oxford: Oxford University Press.
———. (2009) *How We Get Along*. Cambridge: Cambridge University Press.
- Williams, B. (1979) "Internal and external reasons." In *Rational Action*, ed. Harrison, R. Cambridge: Cambridge University Press.