8
Knowledge, Action, and Defeasibility

Carlotta Pavese

8.1. Introduction

One can intentionally do something only if one knows what one is doing while they are doing it. For example, one can intentionally kill one’s neighbor by opening their gas stove overnight only if one knows that the gas is likely to kill the neighbor in their sleep. One can intentionally sabotage the victory of one’s rival by putting sleeping drugs in their drink only if one knows that sleeping drugs will harm the rival’s performance. And so on. In a slogan: Intentional action is action guided by knowledge.¹

This essay reviews some motivations for a ‘knowledge-centered psychology’—a psychology where knowledge enters center stage in an explanation of intentional action (Section 8.2). Then it outlines a novel argument for the claim that knowledge is required for intentional action (Section 8.3) and discusses some of its consequences for the debate about know-how. Section 8.4 argues that a knowledge-centered psychology motivates the intellectualist view that know-how is a species of know-that. In its more extreme form, the view is committed to an epistemologically substantial claim—i.e., that the epistemic profile of know-how is the same as that of propositional knowledge. Now, it is widely believed that know-that can be defeated by undermining and rebutting defeaters (e.g., Chisholm 1966; Goldman 1986; Pollock and Cruz 1999; Bergmann 2000). If that is correct, one corollary of intellectualism is that the defeasibility of know-how patterns with that of knowledge. A knowledge-centered psychology does predict that, for it predicts that both know-how and knowledge are defeated when one’s ability to intentionally act is defeated. In Section 8.5, by replying to a challenge raised in the recent literature (Carter and Navarro 2018), I argue that this prediction is actually borne out.

¹ I am grateful for comments to Adam Carter, Clayton Littlejohn, and Mona Simion.
8.2. Towards a Knowledge-Centered Psychology

A long tradition in the philosophy of mind assigns beliefs a central role in folk psychological explanations of intentional behavior (e.g., Stich 1978; Fodor 1987; Lewis 1976; Stalnaker 1984; Humberstone 1992). More or less explicitly, this tradition confines psychological explanations to an explanation of attempts. Consider the usual example of a psychological explanation, where one’s belief that there is water in the fridge and one’s desire to drink it together are supposed to explain one’s attempt to grab a bottle of water from the fridge. Success happens when the belief is true—when there is indeed water in the fridge. If one’s belief is true, then one will succeed at finding a bottle; if one’s belief is false, one will not succeed at finding water. The dominant thought behind a belief-centered psychology is the idea that, as far as the psychological explanation of behavior goes, whether the world makes the belief true (e.g., whether there is water in the fridge) is irrelevant: what is to be explained is the fact that one attempted to get water from the fridge, whether or not one has succeeded. And one’s belief that there is water in the fridge, together with one’s desire to drink it, suffices to explain one’s attempt, whether or not the belief is true.²

Let us get a bit clearer about the underlying assumption of a belief-centered psychology. Let a condition be something that obtains or fails to obtain at a case, and let a case be a centered possible world or situation—an ordered triple of a subject, a time, and a location. Some conditions are mental, such as that I feel pain or believe that it is raining. Other conditions are non-mental, such as that I have broken my leg. Beliefs are mental conditions and so are attempts—e.g., my trying to get on the bus is a mental condition.³ Like beliefs, attempts are non-factive mental conditions: an attempt to ϕ does not entail successful ϕ-ing, for an attempt might be successful or might fail. As a non-factive condition, an attempt does not encompass those external aspects of the world that make for an agent’s success.

The assumption that psychological explanation should be confined to explaining attempts relies on the idea that actions are decomposable into mental conditions and non-mental conditions—into attempts, on the one hand, and into those external conditions that makes for the agent’s success, on the other. Call this the decomposability assumption and suppose it is true. If so, we can appreciate one of the main motivations that underlie a belief-centered psychology. A psychological explanation of intentional

² See in particular Stich (1978).
³ This terminology follows closely Williamson (2000: Ch. 3).
behavior, understood as attempts, does not need to appeal to anything more than to non-factive mental states—i.e., such as mere beliefs. If when we are explaining behavior, all we are trying to explain is an attempt (e.g., the attempt at finding a bottle of water in the fridge) rather than one’s intentional success (e.g., one’s intentional success at finding the bottle in the fridge) then all we need is a psychological theory that encompasses non-factive attitudes such as beliefs and desires (e.g., one’s belief that water is in the bottle and the desires to drink).

But is the decomposition assumption true? Indeed, there are good reasons to think that actions are not decomposable into mental and non-mental components and that even if attempts were in some sense components of actions, the mentality of actions would not be exhausted by the mentality of attempts. Here is an argument for this conclusion. If attempts exhausted the mentality of actions, then provided that one attempted to $\phi$, one’s eventual success at $\phi$-ing would have to be intentional. For on this picture, the intentionality, and hence the mentality, of an action would be exhausted by its attempt. However, there are a variety of cases in which one attempts at $\phi$-ing, succeeds, and yet fails to act intentionally. That suggests that the intentionality of actions cannot be reduced to the intentionality of attempts; and intentionality being a mark of the mental (e.g., Brentano 1874/1995: 68), that suggests that the mentality of actions cannot be reduced to the mentality of attempts.

A well-known example in action theory which illustrates how attempts and intentionality can come apart is Mary the Bomber (cf. Mele and Moser 1994; Gibbons 2001):

**Mary the Bomber**: Mary intends to kill her uncle by setting off a bomb by a bomb in his house and then, after moving a safe distance away, pressing the large red button on the remote-control device. She does not know much about how these things work and thinks that pressing the button will cause the bomb to detonate but has no idea about the details of this process. Her belief is true and justified and here is what happens. A satellite, launched by the National Security Agency and designed to prevent bombings of just this kind, intercepts Mary’s transmission which causes the satellite to send a warning to the intended victim. But, because of an unfortunate choice of frequency, this causes the bomb to detonate.

Mary killed her uncle and caused the bomb to detonate and did intend both things. But she did not do either of these things intentionally. Hence, the intentionality and mentality of this action is not reducible to the intentionality
and mentality of the attempt. This residual ‘mentality’ of actions calls for an explanation—and presumably an explanation that reduces the intentionality of the action to some mental state of the agent. And for this sort of explanation, belief alone cannot suffice.⁴ For intentional action is a factive mental condition: if one intentionally ϕs, then (trivially) one ϕs. And, if the mental condition to be explained is factive in the way intentional action is factive, then its explanation calls for a factive condition. Non-factive (attempts) might be explainable by non-factive (i.e., beliefs). But factive (intentional actions) ought to be explained by factive (i.e., knowledge).

Mary the Bomber can be accounted for by a knowledge-centered psychology (Gibbons 2001). Mary does not really know that she can provoke the explosion by implementing her plan. That is why her success is too coincidental to count as intentional. More generally, the prediction of a knowledge-centered psychology is that if one’s belief is Gettiered, then one cannot act intentionally on that belief. This prediction is borne out. To see this, consider two more examples. The next is also from Gibbons (2001) but slightly revised:⁵

Cindy and the Lottery: Cindy mistakenly believes that someone rigged a lottery in her favor and that she will be handed the winning ticket at the ticket store. On this basis, she believes of a particular ticket that is being handed to her, that if she buys it, she will win. She buys the ticket and wins. So, her belief that she will win the lottery by buying that ticket is true. It is even justified. Buying a winning ticket is a perfectly reliable way of winning a lottery. Still, intuitively, she did not intentionally win the lottery.

Here again, Cindy intends to win the lottery and attempts to do it by buying a ticket that she believes truly and justifiably to be a winning ticket. This case differs from the previous one, for here no deviant causal chain plays a role in explaining her success (i.e., in making her ticket win). And yet, Cindy’s victory is not intentional: fair lotteries cannot be intentionally won. That suggests, once again, that the intentionality of actions is not exhausted by the intentionality of attempts. But if so, we need a sort of psychology that differs from a belief-centered psychology in that it is not only tailored to explaining attempts.

⁴ Cf. also Levy (2013) and Williamson (2017) for similar remarks about the non-decomposability of actions. Gibbons (2001) also argues that psychological explanations should explain intentional successes and not merely attempts.

⁵ Revised in order to overcome an objection raised by Cath (2015) to Gibbons’ original example. See Pavese (2018) for a detailed discussion of Cath (2015) on Gibbons. Cath’s objection does not extend to Cindy and the Lottery as presented in the text, nor does it extend to the next example from Pavese (2019) — i.e., Daniel and the Barn.
As a final example supporting the claim that intentional action requires more than true belief or even justified true belief, consider the following variation on Carl Ginet’s fake barn case (Pavese 2019; Beddor and Pavese 2019):

**Daniel and the Barn:** along the road to Larissa, Daniel is instructed to stop at any barn that he finds and to await further instructions. However, the road to Larissa passes through fake-barn county. Daniel passes the first barn-looking construction but does not stop for he thinks “There is not enough shadow for me to wait comfortably.” At the second barn-looking construction, Daniel stops and parks. As it turns out, only the second barn-looking construction was a real barn.

In **Daniel and the Barn**, Daniel ends up stopping at a barn and intended to do so. However, intuitively, he did not intentionally stop at a barn. In fact, he would have easily stopped at a fake barn, had the shadow been present there.

In all **Mary the Bomber**, **Cindy and the Lottery**, and **Daniel and the Barn**, the success of the agent does not count as intentional. This fact can be explained on a knowledge-centered psychology. If one possesses knowledge, then one’s belief cannot be lucky (Sosa 1999; Williamson 2000). In all of these examples, the subject’s belief turns out to be true by luck and the same sort of luck undermines the intentionality of their actions. A knowledge requirement on intentional action can explain why the luck of their successes can undermine the intentionality of Mary’s, Daniel’s, and Cindy’s successes: by undermining their knowledge. In contrast, it is unclear that a belief-centered psychology has the resources to explain why these actions are not intentional.

---

6 Cf. also Goldman (1976).

7 Some take fake barn cases to be cases of knowledge (see e.g., Sosa 2007: 31), for they think in those cases, the success of the belief is clearly due to ability. Many, however, disagree about this diagnosis, either on the grounds that it is too little intuitive (e.g., Pritchard 2012; Beddor and Pavese 2018) or on the ground that there is a clear virtue-theoretic rationale for thinking that success in fake barn cases is not attributable to ability. For the latter style of argument, see Littlejohn (2014).

8 Some object to a modal requirement on knowledge. See Beddor and Pavese (2018) for a recent defense.

9 It is important to register that not all luck is epistemically harmful. The kinds of luck that seem most clearly malignant are the ones that seem to show that it is at least partially accidental that a belief ‘turns out’ to be accurate, correct, and so on. When it is lucky that someone believes but not lucky that the belief (given the conditions under which its formed) is correct, the luck does not seem to be malignant. Something similar seems to hold for tryings, attempts, and intentional actions. It seems to matter that the connection between trying and succeeding is lucky or accidental, not whether it’s lucky or accidental that the person tried. Thanks to Clayton Littlejohn for the discussion.
But could one act intentionally on the basis of extremely high probabilistic evidence that however falls short of knowledge? Consider:

**Mary, the Bomber, and the Lottery**: Mary intentionally plays a trillion ticket lottery such that the bomb fails to go off and fails to kill her uncles only if she wins the lottery (and let’s suppose the satellite stuff is left out of the story).

Like in standard lottery cases, arguably Mary does not know that she will lose the lottery (Williamson 2000). But, the objection goes, it seems as though Mary could intentionally kill her uncle by simply intentionally playing that lottery. This is the case even though (like in Gettier cases) she doesn’t know that playing the lottery will detonate the bomb.

In response, granted, Mary does not and cannot know that playing the lottery will detonate the bomb. But there is another thing that she does know—i.e., that it is sufficiently likely that playing the lottery will detonate the bomb, for she knows that it is sufficiently likely that she will lose the lottery. In fact, as I will explain later in the essay, it is independently plausible that probabilistic knowledge of this sort is central to an explanation of intentional action (Pavese 2020). To see this, consider Davidson’s (1971: 50) example:

**Carbon Copying**: I have a stack of carbon paper in front of me. In order to save time, I try to sign the top page of the stack with enough force to that I sign all the copies simultaneously. I do this despite the fact that I do not believe that I will succeed. In this case, I might succeed in this endeavor while failing to know that my signature would be legible on the last page because I do not have sufficient confidence for full belief.

This seems to be a case of intentional action. But it is tempting to say in this case that I do not have knowledge because of my lack of confidence. This would be too quick, however. As several people have argued in the recent literature (Weisberg 2013; Moss 2018), credences can amount to knowledge too. Although I do not know that I will succeed at signing all the carbon copies, I still estimate that I will succeed with some probability. And although I do not know that I will succeed, I still know that it is sufficiently likely that I will. And, knowing that requires believing that it is sufficiently likely that I will. This belief can be modeled as a credence of sufficiently high degree that I will sign all the carbon copies. If that is correct, then Davidson’s (1971) carbon copy

---

10 Thanks to Adam Carter for the example.
example—or Mary, the Bomber, and the Lottery, or other similar examples\(^{11}\)—raise no issue for a knowledge requirement on intentional action, for what guides intentional action in these cases can still be knowledge, albeit probabilistic.

In conclusion, explanations of intentional action do seem to essentially appeal to knowledge. These explanations are plausibly psychological in nature, for they aim at explaining a factive mental condition (intentional action) in terms of a more primitive factive mental condition (the mental state of knowledge).\(^{12}\) As such, they provide support to a knowledge-centered psychology, for they support the view that knowledge plays a central role in psychological explanations of intentional action.

### 8.3. Intentional Action does Require Knowledge

The claim that intentional action requires knowledge has been recently challenged. It is instructive to consider the challenge for it points towards a more principled argument for thinking that knowledge is required for intentional action.

Cath (2015: 11) argues that one can have intentional action without knowledge, upon considering cases like the following:

**Bob the Pilot**: Bob wants to learn how to fly in a flight simulator. He is instructed by Henry. Unbeknownst to Bob, Henry is a malicious imposter who has inserted a randomizing device in the simulator’s controls and intends to give all kinds of incorrect advice. Fortunately, by sheer chance, the randomizing device causes exactly the same results in the simulator as would have occurred without it, and by incompetence Henry gives exactly the same advice as a proper instructor would have done. Bob passes the course with flying colors. He has still not flown a real plane. Bob has a

\(^{11}\) Setiya (2012) puts forward an example of the subject believing their hand to be paralyzed who tries nonetheless to clench their fingers. This case too can be handled by appealing to probabilistic knowledge, for although one does not know that one will succeed by intending to succeed, one knows that one might succeed by intending. Cf. Pavese (2020). For a different sort of diagnosis of Setiya’s case, cf. Pavese (2018).

\(^{12}\) Gibbons (2001) emphasizes that the role of knowledge in explaining intentional action provides a novel argument for the claim that knowledge is a mental state—a claim famously defended by Williamson (2000) and Nagel (2013). Although both Williamson (2000) and Nagel (2013) provide arguments for the role of knowledge in explaining behavior, neither focuses specifically on the role of knowledge in action theory. For a comparison, see Pavese (2019).
justified true belief about how to fly, but that justified true belief does not amount to knowledge.¹³

Cath uses this example to argue that one can have know-how even though one’s relevant belief is Gettiered. Cath claims that intuitively Bob can intentionally fly. And for some action \( \phi \), if one can intentionally \( \phi \), then one knows how to \( \phi \) (Williamson and Stanley 2001; Hawley 2003; Setiya 2012; Cath 2015; Pavese 2018, 2020). Therefore, Bob must know how to fly, even though he does not know the relevant instructions.

Cath thinks that this diagnosis is supported by comparing the case of Bob with the case of Joe, who is a near perfect counterpart of Bob except that his belief is not Gettiered: his simulator operated correctly and did so non-accidentally; his instructor intentionally gave him the correct advice, and so on. If Joe were to try to fly a plane in normal circumstances, he would typically succeed in so doing and his successful actions would be unquestionably intentional actions. And it is an implicit stipulation of the flight simulator case that if Bob were to try to fly a plane in normal circumstances then he would be just as likely to succeed as Joe. Cath contends that not only would Bob succeed as often as Joe but, like Joe, his actions would appear to have all the standard kind of properties that are thought to distinguish merely successful actions from intentional actions. Hence, according to Cath, it is plausible that Bob’s successful actions of flying, like Joe’s, would be perfectly under his control or guidance as he performs them. But is Cath (2015) right in assuming that, in Bob the Pilot, Bob’s performance is under his control?

Both intuitive and theoretical considerations suggest that the answer to this question ought to be “No.” Compare Joe and Bob. Note that, strictly speaking, Joe knows what he is doing while he flies the plane: he knows that he is following instructions that are conducive to successful flying. In contrast, by assumption, Bob does not know that. Hence, he does not know what he is doing while he flies the plane. But consider how unintuitive it is to ascribe intentional action to one who lacks knowledge of what they are doing:

**Awful**: Bob intentionally landed the plane, but he did not know that he was landing the plane by following the given instructions.

¹³ This example is initially due to Stanley and Williamson (2001) who actually give it as an example showing that know-how is incompatible with epistemic luck. Poston (2009) also discusses it and so does Stanley (2001: Ch. 8).
A similar intuition has been voiced by several philosophers in the past. Hampshire claims that “[I]f a man is doing something without knowing that he is doing it, then it must be true that he is not doing it intentionally” (Hampshire 1959: 95). Anscombe (1959: Section 8.8) holds that if someone is φ-ing intentionally, she knows without observation that she is φ-ing. More recently, Gibbons claims that “talk of intentional action presupposes a certain degree of control on the part of the agent. Control, like perception, requires the right kind of connection between the agent and the facts. An essential ingredient in this kind of connection is knowledge” (2001: 591).

Hence, the intuition that an intentional action’s control depends on one’s knowledge is very widespread. Of course, no important philosophical claim should be motivated merely by intuitions, for they are too unstable and possibly theoretically driven to be conclusive. There is, however, a positive theoretical argument to the effect that, contra Cath, one’s actions cannot be intentional unless they are guided by knowledge.¹⁴

Here is the gist of the argument. Cath assumes intentional action ought to be “under one’s control” and assumes that Bob’s success is under his control. But as in Mary the Bomber, the belief in question is Gettiered, and so happens to be true by luck. According to the standard account, a belief is lucky just in case it could easily fail to be true. If so, both Bob and Mary could easily have the beliefs they have even if they were false. In this case, an attempt of theirs based on those beliefs could easily fail. Arguably, then, their success is too lucky to count under their control. More precisely, here is the argument step by step:

(a) If one’s action is intentional, then it is under one’s control. (Premise)
(b) If one’s action is (too) lucky, then it is not under one’s control. (Premise)
(c) If an action is based on a Gettiered belief about how to perform it, then it is (too) lucky. (Premise)
(d) Hence, if an action is based on a Gettiered belief about how to perform it, then it is not under one’s control. (From b, c)
(e) Hence, if an action is based on a Gettiered belief about how to perform it, then that action cannot be intentional. (From a, d)

¹⁴ Greco (2016) considers a different sort of argument for the claim that knowledge is necessary for explaining action—one that relies on the nature of explanation as counterfactually robust. See also Pavese (2018).
As rarely happens in philosophy, we have here a deductively valid argument to the effect that intentional actions ought to be based on knowledge. Is the argument sound?

Premise (a) relies on rather minimal assumptions about the nature of the action, widely endorsed in action theory (e.g., Mele and Moser 1994; Gibbons 2001), according to which intentional action is under the agent’s control. (Because this premise is actually granted by Cath (2015: 10–11), I will assume it without much argument.)

Premise (b) is the rather uncontroversial claim that, if a certain success is (too) lucky, then it is not under the control of the agent. This is a common assumption in action theory, where it is customary to infer from the fact that an act is too coincidental that it is, therefore, not intentional (e.g., Melse and Moser 1994: 40), as well as in debates on moral responsibility (Nagel 1979: 59; Williams 1981: 126, 1993), where moral luck is deemed to be incompatible with control.¹⁵⁶

The most controversial premise is, I take it, Premise (c)—i.e., the claim that, if an action is based on a Gettiered belief, then the action is (too) lucky to qualify as under the agent’s control. Here is an argument for thinking that it is true. The following is a plausible sufficient condition on a lucky action:

**Sufficient**: if S succeeds at φ-ing at a world @ but S fails to φ in many of the sufficiently close worlds where S tries to φ, then S’s φ-ing at @ is (too) lucky.¹⁷

Note that **Sufficient** does not require that for one’s action at @ to be not-too-lucky, one succeed at performing it when one tries it in every sufficiently close world.¹⁸ This would be too demanding a requirement: as Austin (1961: 218) put it, “a human ability or power or capacity is inherently liable not to produce success, on occasion, and that for no reason.” If Austin is right, many actions might be under the control of their agent and even manifest their skills and

---

¹⁵ It is worth noting that some recent work takes moral luck to be a species of a larger genus of luck, of which there are other species as well, such as *epistemic luck*. Such an approach does not build in the idea that luck is opposed to control. See Pritchard (2006) and Coffman (2015) for similar approaches. Thanks to Mona Simion for the discussion.

¹⁶ It is, of course, a difficult and partly a context-sensitive matter, what counts as “under one’s control.” But, for our purposes, we do not need to settle this question, the idea being that if a certain success counts as too lucky, it is not sufficiently under the control of the agent to count as intentional.

¹⁷ Beddor and Pavese (2018: 6) for a defense of this sort of condition on lucky actions.

¹⁸ Although strictly speaking only attempts might fail to succeed, it will simplify my exposition to talk as if one’s action might fail to succeed in some worlds where it is tried.
still fail in some close worlds, for no particular reason. Rather, according to Sufﬁcient, an action is too lucky if it fails in more than just some close worlds when tried—if it fails in many close worlds (for that is what “many” means: more than just some).

My argument for Premise (c) is that actions based on Gettiered beliefs are doomed to satisfy Sufﬁcient, and hence are doomed to be too lucky in the relevant sense. Here is why. If an action is based on a true belief about how to bring it about, then it might nonetheless fail when tried in some close worlds. In fact, even if the action is based on knowledge, it might still fail in some close worlds. That might be because in those close worlds, something interferes with the agent’s basing their action on the relevant belief. For example, an archer’s shot might fail in a counterfactual circumstance where the archer gets distracted by a passer-by and hence fails to appropriately base their shot on their knowledge about how to shoot. Or it might fail also because, although their shot is appropriately based on their knowledge, the world does not cooperate in the way required for success. For example, the archer’s shot might fail in a circumstance where a fluke happens, and an unexpected gust of wind interferes. Or it might fail because their hand was for no reason slightly less ﬁrm than they expect it to be a moment ago. In those counterfactual circumstances, the belief on which the agent’s performance is based might still be true—it is still true that one must do this and this to shoot the target under certain conditions (e.g., in non-windy conditions and when one’s hands are ﬁrm). It is just that there the world does not cooperate in the way required for success.

Now, when an agent acts on a Gettiered belief, there will be an additional reason for why they might fail in some close worlds—i.e., because the belief they act upon there is false and so does not accurately represent the world. Gettiered beliefs are generally assumed to be lucky precisely in the sense that they might fail to be true in some close worlds (Sosa 1999; Williamson 2000; Pritchard 2005). If so, then an action based on a Gettiered belief will fail in more close worlds than it would have if it had been based on knowledge. For it will not just fail in those close worlds where it is not properly based on the relevant belief or where, even though properly based, the world does not cooperate. It will also fail in those close counterfactual circumstances where the basing belief about how to perform the action is false. Hence, it will fail in more than just some close worlds. Hence, an action based on a Gettiered belief is doomed to satisfy Sufﬁcient. And if it satisﬁes Sufﬁcient, then Premise (c) is true.

¹⁹ Thanks to Clayton Littlejohn for discussion on this point.
This concludes my argument for Premise (c). With Premise (a) to Premise (c) in play, the conclusion deductively follows actions based on Gettiered beliefs cannot be intentional. Intentional action ought to be guided by knowledge.

8.4. From a Knowledge-Centered Psychology to Intellectualism About Know-How

Hence, a knowledge-centered psychology can be motivated on the basis of intuitions about cases (Section 8.2). But there are, also, more theoretical considerations on behalf of a knowledge-centered psychology, ones that rely on rather minimal assumptions about the nature of Gettiered beliefs, control, and intentional actions (Section 8.3). This section explores some of the consequences of a knowledge-centered psychology and details an argument from a knowledge-centered psychology to an intellectualist view about know-how.

First, note that the kind of knowledge that, on a knowledge-centered psychology, explains intentional action is exactly the same kind of knowledge that, on a broadly intellectualist picture, is required by know-how. To see this consider the kind of knowledge that would be needed to explain intentional action. Start with Goldman’s (1970) action theory, according to which one intentionally \( \phi \) when one has a plan to \( \phi \), where a plan to \( \phi \) is a belief that specifies the means to \( \phi \) (cf. also e.g., Audi 1986; Bratman 1987; Ginet 1990; Harman 1976; Velleman 1989, 2007; Mele and Moser 1994):

\[ \text{(Intentionality/Belief): if } s \text{ intentionally } \phi \text{, then there are some means } m_1, \ldots, m_n \text{ to } \phi \text{ such that } s \text{ truly believes that } m_1, \ldots, m_n \text{ are means for them to } \phi. \]

\[ \text{(Intentionality/Knowledge): if } s \text{ intentionally } \phi \text{, then there are some means } m_1, \ldots, m_n \text{ to } \phi \text{ such that } s \text{ knows that } m_1, \ldots, m_n \text{ are means for them to } \phi. \]

It is independently plausible that the content of one’s knowledge in \( \text{(Intentionality/Knowledge)} \) ought to be spelled out in probabilistic terms. To see this, start with \( \text{(Intentionality/Knowledge)} \). It requires of intentionally \( \phi \)-ing that one have knowledge, of some means, that they are means for one
to ϕ. What does it mean for some means to be means for one to ϕ? Not that, for some way ψ of ϕ-ing, one will ϕ by ψ-ing: that is far too strong, for one might intentionally ϕ even though one has some doubts about whether one will succeed (Goldman 1970; Harman 1976). Recall, for example, Davidson’s (1971: 50; 1980: 91–4) carbon copy example, discussed earlier.

Should the relevant knowledge be that, for some means ψ of ϕ-ing, one would in most cases succeed at ϕ-ing by ψ-ing? This is also too strong: one might intentionally ϕ even though one might fail in most circumstances, as the baseball player who fails at batting nineteen times out of twenty may nonetheless intentionally bat the one time they succeed. That suggests that the relevant knowledge is that, for some means ψ of ϕ-ing, one could ϕ by ψ-ing. But what does it mean that one could ϕ by ψ-ing, if not that one is sufficiently likely to ϕ by ψ-ing, where what counts as sufficiently likely may vary from task to task? This gives us:

(Intentionality/Probabilistic Knowledge): if s successfully and intentionally ϕs at t, then at t s knows, for some means ψ of ϕ-ing, that oneself is sufficiently likely to ϕ by ψ-ing.

Now, according to standard formulations of intellectualism (Stanley and Williamson 2001; Stanley 2011; Pavese 2015, 2017), one knows how to ϕ only if, for some means ψ to ϕ, one knows that ψ is a means for one to ϕ:

(Intellectualism about Know-How) s’s knowing how to ϕ is at least in part of a matter of knowing, for some means ψ to ϕ, s knows that ψ is a means for them to ϕ.²

Consider Intellectualism about Know-How. In the original formulation, it is the view according to which knowing how to ϕ is at least in part a matter of knowing that certain means are means for one to ϕ. But what does that mean? We do not want to require, for some means to be means for one to ϕ that one’s ψ-ing invariably result in one’s successfully ϕ-ing; nor that it result in one’s successfully ϕ-ing in most cases. That would be too demanding: after all, Babe Ruth does know how to hit a home run and yet fails at successfully hitting a home run in many circumstances. In order for ψ to be a way for one

² I am stating Intellectualism as the view that know-how requires knowledge of the means for action, rather than as a fully reductive claim. Stating Intellectualism as a fully reductive claim would require talking about practical modes of presentation, which I cannot discuss here. See Pavese (2015, 2017, 2019a).
to $\phi$, all that is required is that one be sufficiently likely to successfully $\phi$ by $\psi$-ing, where what counts as “sufficiently likely” may vary with the task at hand (and the circumstances under which the task is being performed). This gets us to:

\begin{quote}
\textit{(Intellectualism about Know-How):} $s$ knows how to $\phi$ only if, for some means $\psi$ for $s$ to $\phi$, $s$ knows that it is sufficiently likely for \textit{them} to succeed at $\phi$-ing by $\psi$-ing.
\end{quote}

Hence, if one unpacks the clause \textit{means for one to $\phi$}, a plausible upshot is that both \textit{(Intellectualism about Know-How)} and \textit{(Intentionality/Knowledge)} should be stated through a \textit{probabilistic language}: they both require knowledge that one is sufficiently likely to succeed at $\phi$-ing through certain means. Following Moss (2018), one might know that it is sufficiently likely for oneself to succeed at $\phi$-ing by $\psi$-ing in virtue of possessing a sufficiently high credence that one will succeed at $\phi$-ing by $\psi$-ing. Because possessing this credence does not require that one grasp the concepts of likelihood or probability, Pavese (2020) argues that this rendition of \textit{(Intellectualism about Know-How)} overcomes the challenge of \textit{over-intellectualization} that Setiya (2012) and other authors have raised against it.

Hence, the knowledge that \textit{(Intentionality/Knowledge)} requires for intentional action is the same that intellectualists require for know-how. I suggested that we can use a knowledge-centered psychology to argue for intellectualism about know-how. How would such an argument go? Start from \textit{(Know-How/Intentionality)}, endorsed by intellectualists and anti-intellectualists alike (Ryle 1949; Stanley and Williamson 2001; Stanley 2011; Hawley 2003; Hornsby 2004, 2011; Setiya 2012; Pavese 2018):

\begin{quote}
\textit{(Know-How/Intentionality):} if $s$ intentionally $\phi$s, $s$ knows how to $\phi$.
\end{quote}

Among the motivations behind \textit{(Know-How/Intentionality)} is the idea that operations which cannot be performed intentionally, such as digesting, are ones that one cannot know how to perform (Williamson and Stanley 2001). Moreover, manifestations of know-how seem to be characteristically intentional: as Ryle (1949) put it, what distinguishes the clumsy person, who falls and tumbles by accident, and the skillful clown is that the latter, but not the former, falls and tumbles on purpose.

Further, suppose that \textit{(Intentionality/Knowledge)} is true and so that the intentionality of an action is to be explained at least in part in terms of propositional
knowledge. Then by \((\text{Know-how/Intentionality})\) and \((\text{Intentionality/Knowledge})\), we get that, if one intentionally \(\phi\)s, one both knows how to \(\phi\) and has propositional knowledge of some means to \(\phi\):

\((\text{Know-how, Intentionality, Knowledge})\): if \(s\) intentionally \(\phi\)s, \(s\) both knows how to \(\phi\) and for some means \(m_1, \ldots, m_n\), knows that means \(m_1, \ldots, m_n\) are means for then to \(\phi\).

The intellectualist picture provides the best explanation for why \((\text{Know-How, Intentionality, Knowledge})\) should hold. According to this explanation, \((\text{Know-How, Intentionality, Knowledge})\) is true not just out of a coincidental aligning of propositional knowledge and know-how in intentional action. Rather, its truth is grounded in the very nature of know-how.

By mostly appealing to a linguistic argument when motivating their views (Stanley and Williamson 2001; Stanley 2011), intellectualists have sold the view short.²¹ The chief motivation for the view does not come from linguistics: it comes from the sort of action theory that a knowledge-centered psychology recommends.

### 8.5. The Defeasibility of Know-How

I have argued on behalf of a knowledge-centered psychology by looking at the role of knowledge in explaining intentional action. Knowledge-centered psychology naturally goes together with an intellectual picture of know-how that vindicates the relation between know-how and intentional action.

Now, the intellectualist picture motivated by a knowledge-centered psychology and outlined in the previous section makes a very clear prediction: that know-how is defeated exactly when knowledge is defeated. Against this prediction, Carter and Navarro (2018) argue that the defeasibility of know-how does not go together with the defeasibility of knowledge. They use this claim to argue against the intellectualist claim that know-how consists in a state of propositional knowledge. We are now in a position to assess the problems with Carter and Navarro’s (2018) argument, which turns on a failure to appreciate the relation between knowledge and intentional action.

²¹ For worries concerning the linguistic argument on behalf of intellectualism, see Brown (2013).
Carter and Navarro (2018: 666) propose the following example:

**Ana and the Grenade Factory**: Ana and María work in a grenade factory during the Spanish Civil War. They are thoroughly instructed when hired, with examples and practical explanations. By controlled trial and error, they learn their job, and both continue working at the factory for years, believing they are making working grenades. However, one day each comes to realize that the other is making grenades in an importantly different way, and they identify the origin of the problem: as it turns out, the instructions were ambiguous and allowed for two different interpretations. The instructors were not aware of this, and there is nobody above them now who may say who is right. Given that the grenades may only be used in battle, which is very far away, neither Ana nor María knows whose grenades actually work, and so there is no way to find out who is making them the right way. As a matter of fact, Ana got the instructions right (she produces grenades in way w, which is the correct way); she is very successful in producing grenades that later work perfectly. It is María who got something wrong (she makes them in w’, the possible interpretation of the instructions that the instructors did not foresee), and her grenades are always duds. Unaware of this, both have reasonable doubts they did not have before, but they have to keep on working.

According to Carter and Navarro (2018), before receiving information about how her knowledge has been acquired (call this piece of information ‘MISLEADING’), Ana might know, for some means to make grenades, that it is a means to make grenades; but her knowledge is defeated as soon as the misleading evidence is acquired. On the other hand, they think that Ana still knows how to make grenades after receiving MISLEADING. If they were right, this would be a case where know-how stands undefeated whereas the corresponding knowledge is instead defeated. They conclude (2018: 669):

If know-how really were a case of know-that, we should expect it to be defeasible by the same kinds of mechanisms by which propositional knowledge is defeated. But it is not. In other words: garden variety defeaters of knowledge-that do defeat the knowledge agents have about the ways in which they do what they do.²²

²² The argument assumes that knowledge can be easily defeated by high-order evidence. This assumption is controversial and is not granted by prominent epistemologists (Aarnio 2010, 2014). Let us play along, however, and see that the putative challenge rests on other false assumptions about the nature of know-how.
Carter and Navarro’s argument hinges on two claims. The first claim is that the relevant propositional knowledge is defeated in this case; the second is that Ana still knows how to make grenades, upon receiving MISLEADING. Let me grant the former claim and focus on the latter.

Why think Ana still knows how to make grenades, upon receiving MISLEADING? The intuition that she does is not nearly as robust as they seem to think. And the only argument Carter and Navarro (2018: 666) provide in support of their intuition is the following:

The claim that Ana preserves her know-how along all the variations of the case is supported by the fact that she is still able to make grenades proficiently, and the doubts she acquires do not seem to imperil this ability in any relevant sense.

Carter and Navarro (2018) are going far too quickly here. Granted, Ana still preserves some ability that is relevant to grenades-making. What is much less obvious, and as I argue ultimately incorrect, is to assume that Ana preserves the sort of ability that goes together with know-how.

As the discussion in the previous section already suggests, know-how does not just go together with any ability. It goes together with the ability to intentionally perform a certain task. For example, knowing how to make risotto does not merely go together with the ability to make risotto but with the ability to intentionally make risotto. For if one had the ability to make risotto but lacked the ability to intentionally make it, one would not count as knowing how to make risotto. This point is well-known in the literature at least since Hawley (2003) and is accepted by both intellectualists and anti-intellectualists (Ryle 1949; Setiya 2012; Pavese 2017). For example, the clumsy person has the ability to fall and tumble, as they reliably do so. But only the clown has the ability to intentionally do that. As another illustrative example, Susie may have the ability to irritate Ben, for she would succeed at irritating him if she tried. But suppose she falsely believes that it is the smell of the smoke, rather than the noise she makes whenever she smokes, that irritates Ben. In this case, she does not intentionally irritate Ben: her success is too coincidental to count as intentional. Because of this, it seems that she does not know how to irritate Ben. On the bases of similar examples, intellectualists and anti-intellectualists alike endorse the claim that know-how goes together with the ability to intentionally perform the task.

Suppose it is true that know-how goes together with the ability to intentionally perform a task. If so, it is independently plausible that there is an important
sense in which, upon receiving MISLEADING, Ana does not know any longer how to make grenades. For upon receiving MISLEADING, when asked to make grenades, she will be at a loss. Not only that: she will also refuse to teach others how to make grenades. And she will even stop performing at the workplace, until she is told that she has been making grenades correctly all along. Suppose she were forced to reproduce whatever process she initiated before MISLEADING. She would unknowingly succeed at making grenades. But the success would be too out of her control to count as intentional. She is still able in some sense to make grenades but in an important sense she now lacks the ability to intentionally make grenades. If so, then she also lacks know-how.

If the reader is not yet ready to grant this conclusion, it is because, actually, things are more complex, and some additional distinctions are called for. Ascriptions of abilities of the form “s can intentionally perform a task” are opaque, for as it is well known in action theory, “intentionally” is an intentional operator (cf. Davidson 1971; Goldman 1970). For example, Lois might intentionally kiss Superman but not intentionally kiss Clark Kent. Because of the opacity of intentionality reports, it is paramount to distinguish between (de re ability) and (de dicto ability):

\[(\text{de re/ability}): \text{there is some task } t \text{ that is in fact the task of making grenades such that Ann has the ability to intentionally perform } t.\]

\[(\text{de dicto/ability}) \text{ Ana has the ability to intentionally make grenades.}\]

While (de re/ability) ascribes Ana a de re ability, (de dicto/ability) ascribes Ana a de dicto ability.

Now, with this distinction in play, consider again Ana’s situation upon receiving MISLEADING. (De dicto/ability) is now false: Ana does not have a de dicto ability any longer. For one to have the relevant de dicto ability, one needs to be able to make grenades on demand (to be in a situation such that, if asked to make grenades, Ana would do so). Ana does not have that ability: were she asked, after receiving MISLEADING, to make grenades, she would now be at a loss. If know-how goes with the ability to intentionally perform a task, then, to this distinction between a de dicto ability and a de re ability, there corresponds the distinction between de re know-how and de dicto know-how:

\[(\text{de re/KH}): \text{there is a task } t \text{ that happens, unknown to Ana, to be the task of making grenades, such that Ana still knows how to perform } t.\]

\[(\text{de dicto/KH}): \text{Ana knows how to perform grenades.}\]
(de re/K) and (de dicto/K) ascribe different kinds of know-hows—de re know-how and de dicto know-how—which go along with different dispositions in behavior. One might have de re know-how even if one has no idea that what one is doing when doing t is making grenades. Suppose, for example, one is simply instructed to follow a certain procedure but has no idea of its outcome. In this case, one might have de re know-how without de dicto know-how. This is plausibly Ana’s quandary: Because Ana still knows how to execute whatever task she was executing before MISLEADING, she plausibly still has de re know-how. After all, if she were told at the workplace to do whatever she was doing before she received MISLEADING, and she obeyed the order, she would intentionally perform a task, which, unknown to her, is the task of making grenades. Hence, Ana preserves her de re ability upon receiving MISLEADING. So, Ana plausibly also retains her de re know-how upon receiving MISLEADING.

However, Anna does lose de dicto know-how. For her to possess de dicto know-how, it is not sufficient to possess de re ability; she would need in addition to have the corresponding de dicto ability, which as we have seen she lacks. While Ana loses de dicto know-how and de dicto ability upon receiving MISLEADING, Ana preserves de re know-how and de re ability, for she still knows how to do whatever it was that she was doing before (which, as far as she knows, is not accurately making grenades!), and she still preserves the de re ability to make grenades upon receiving the misleading information.

Crucially, intellectualists can accept all of this. According to intellectualism, de re know-how only requires de re knowledge while de dicto know-how requires de dicto knowledge:

(de re/K): there is a task t that is in fact, but unknown to Ana, the task of making grenades such that Ana knows for some way w that w is a way to execute t.

(de dicto/K): Ana knows for some way w that w is a way to make grenades.

Upon receiving MISLEADING, (de dicto/K) becomes false: Ana loses the relevant de dicto knowledge. However, Ana arguably still also preserves the relevant de re knowledge (ascribed by (de re/K)). MISLEADING only defeats (if anything) her de dicto knowledge—i.e., the knowledge that the procedure she was implementing was for making grenades. Her de re knowledge, instead, is not at all defeated by MISLEADING: Ana continues to have it, as she might continue to know what procedure she was following before receiving...
MISLEADING, when she was intending to make grenades. And so, by intellectualism’s lights, she might continue to know how to make whatever she was making when she thought (correctly, it turns out!) that she was making grenades: because Ana retains her de re knowledge, by intellectualism’s lights, Ana can retain her de re know-how as well as her de re ability.

Let me end by considering two possible responses.

I have argued that Ana might lose her de dicto know-how upon receiving MISLEADING, while possibly retaining her de re know-how. Could not Carter and Navarro (2018) reply that Ana does retain her de dicto know-how and her de dicto ability all along but upon receiving MISLEADING, she simply cannot act on those, because of her new doubts? Compare: many Olympic gymnasts know how to do the fancy tricks they do even though many—in the heat of the competition—have doubts about whether they can do them successfully.

In response, the analogy with the Olympic gymnasts is misleading. In the case of Olympic gymnasts, it is plausible that despite their doubts, they still know how to perform their fancy tricks (de dicto). After all, they can still intentionally do their fancy tricks (de dicto) outside of the heat of the competition. Their ability to intentionally act on that knowledge is not lost but only ‘masked.’ This can be explained on the current picture: These athletes retain their (de dicto) knowledge all along and simply cannot access it in some circumstances. By contrast, Ana has lost her de dicto knowledge and unless she regains it, there is no circumstance where she can still intentionally make grenades (de dicto). In Ana’s case, then it is utterly implausible that her ability to intentionally act is simply masked.

A second possible response goes as follows. Maybe, Ana does preserve her de dicto know-how and her de dicto ability upon receiving MISLEADING. What she lacks is knowledge that she does know how and that is what explains the lack of de dicto ability. If one embraces this position, one commits oneself to replacing (Know-how/Intentionality)—a claim that, as we have seen, both intellectualists and anti-intellectualists agree upon—with a considerably stronger claim that intentionally ϕ-ing requires knowledge that one knows how to ϕ. This stronger claim is rather implausible. For one thing, some non-human animals can certainly intentionally act, though lacking the concept of know-how. For another, suppose one learns how to make grenades but has not had the occasion to form a belief one way or another about whether one has indeed learned. If one is asked to produce one and one tries, surely one can still intentionally do it (even de dicto!), even though one does not know that they knew how to make them.
8.6. Conclusions

There is a lot going for a knowledge-centered psychology: it explains our intuitions in a variety of cases where intentionality of an action seems absent because of the agent’s lack of knowledge (Section 8.2). The role of knowledge in explaining intentional action is also demonstrated by more theoretical considerations showing that an action cannot be under one’s control unless it is guided by knowledge (Section 8.3). A knowledge centered psychology, in turn, motivates intellectualism about know-how, for it explains why know-how and the ability to intentionally act go hand in hand (Section 8.4).

Having motivated a knowledge-centered psychology, I have appealed to it in a discussion of Carter and Navarro’s (2018) argument to the effect that know-how differs from knowledge in its pattern of defeasibility (Section 8.5). I have argued that Carter and Navarro’s (2018) challenge fails, for they fail to show that know-how remains undefeated when knowledge is defeated. Their alleged challenge turns on the failure to appreciate the relation between knowledge, know-how, and intentional action. Because of that, they fail to distinguish between different sorts of abilities that go together with know-how. Once one appreciates that know-how goes with the ability to act intentionally, because ascriptions of this sort are opaque, it becomes paramount to distinguish between de re abilities and de dicto abilities. With this distinction comes the corresponding distinction between different sorts of know-hows and between the different sorts of knowledge that Ana preserves or loses upon receiving MISLEADING. As we have seen, against Carter and Navarro (2018), the sort of de re abilities Ana does preserve can be fully accounted for on a picture on which know-how is knowledge. And those de dicto abilities that she does lose are also correctly predicted to get lost on the same intellectualist picture.

Far from coming apart in their pattern of defeasibility, know-how and knowledge go hand in hand, just as one would expect on the sort of intellectualist picture that is motivated by a knowledge-centered psychology.

References


