

Knowledge and Mentality *

Carlotta Pavese

forthcoming in *Philosophical Perspectives*

March 8 2021

Abstract

This paper reexamines the case for mentality — the thesis that knowledge is a mental state in its own right, and not only derivatively, simply by virtue of being composed out of mental states or by virtue of being a property of mental states — and explores a novel argument for it. I argue that certain property singled out by psychologists and philosophers of cognitive science as distinctive of skillful behavior (*agentive control*) is best understood in terms of knowledge. While psychological theories of agentive control that appeal to monitoring mechanisms, such as attention, have been proposed, these theories cannot account for the full scope of controlled action. By contrast, I argue that an epistemic theory of agentive control that invokes knowledge is extensionally adequate. It is when it comes to understanding the hallmarks of skillful performance that the theoretical benefits of thinking of knowledge as mental can be fully appreciated.

1 Introduction

Is knowledge a mental state? While some have argued for a positive response to this question (Prichard (1950), Guttenplan (1994), McDowell (1995), Peacocke (1999), Williamson (2000), Nagel (2011), Gibbons (2001)), most philosophers respond negatively to it (e.g., Stich (1978),

*I am grateful to Yair Levy, Timothy Williamson, and Sarah Moss for discussions on the topic of this paper. The epistemic theory of control that is outlined in §7 emerged from some joint work with Bob Beddor (Beddor & Pavese (2020), Beddor & Pavese (2021)), though I take full responsibility for the particular version defended here and for the arguments that I have put forward for it in this paper.

Jackson (2002), Magnus & Cohen (2003), Sosa (2009b), Smith (2017), Fricker (2009), McGlynn (2014), McGlynn (2017), Papineau (2019)). The ‘mentalists’ think that knowledge is a mental state in its own right, in the same way that desires, beliefs, fears, and intentions are considered to be mental states; by contrast, the ‘non-mentalists’ take knowledge to be mental only derivatively, if at all, by virtue of being composed of a mental state, or by virtue of being a property of a mental state of belief. This latter group of philosophers converge in denying MENTALITY:

MENTALITY Knowledge is a mental state in its own right and not only derivatively, simply by virtue of being composed of a mental state, or by virtue of being a property of some mental state.

According to MENTALITY, the mind is *world-encompassing* in the sense that whether or not certain propositions are true can be read off from the complete specification of a subject’s mental states that have those propositions as their contents. For example, according to MENTALITY, if one knows which team has won the Italian soccer league this year, and Inter Milan won it, then that Inter Milan won the Italian soccer league this year supervenes on one’s mental state.

In this paper, I reexamine the case for MENTALITY and explore an argument for it that emphasizes the role of knowledge in a theory of agentic control. I suggest that *agentic control* — uniformly singled out by psychologists and philosophers of cognitive science as distinctive of skillful behavior — is best understood in terms of knowledge. If this is correct, a theory of agentic control ought to take the mind to be world-encompassing. I argue that psychological theories of agentic control in terms of monitoring mechanisms, such as focus and attention, cannot account for the full scope of controlled action but that a theory of agentic control that invokes knowledge can. I argue that my case for MENTALITY improves on extant ones in some important respects.

Here is the plan. §2 reviews the two most prominent cases for MENTALITY and discusses possible reasons for discontent. This will set the stage for my own argument. §3 introduces a discussion of agentic control in the psychology of skills as well as in philosophical discussions in action theory. §4 argues that any attempts to understand agentic control in terms of monitoring mechanisms, such as attention and focus, are bound to fail. §5 discusses and rebuts some reasons for skepticism regarding the attempt to analyze controlled and intentional actions in terms of more basic conditions — the ‘intentional action first program’. In §6, I argue that agentic control is prime — in Williamson (2000)’s sense of being non-decomposable into internal and external conditions — and that this suggests that it can only be reduced to other prime conditions. §7

provides an argument for an epistemic theory of control, which combines Anscombe (1963)'s requirement for intentional action — practical knowledge — and a knowledge-how condition, understood as knowledge of means for performing an action. §8 compares my argument to other arguments for MENTALITY. §9 concludes.

2 Williamson and Nagel on Mentality

2.1 The Case for Mentality in *Knowledge and Its Limits*

In *Knowledge and Its Limits*, Williamson provides the first systematic case for MENTALITY. While the argument offered is complex and articulated, the general strategy is *defensive*. Williamson starts his argument by motivating the claim that MENTALITY should be the *default view*: “our initial presumption should be that knowing is a mental state” (Williamson, 2000, pp.22-23). Williamson points out that, prior to theory building, we have learned the concept of the mental by examples, which include believing and desiring as well as knowing. Moreover, pretheoretically, there do not seem to be important differences between factive states and non-factive states that disqualify the former from being mental. Indeed, that the mental includes knowing and other factive attitudes seems to be built into the ‘natural understanding of the procedure by which the concept of the mental is acquired’. So, according to Williamson, while there may be reasons to dismiss those pretheoretical considerations on the basis of other more pressing theoretical reasons, those theoretical reasons should undergo close scrutiny. Therefore, in the rest of his book, Williamson goes on to show that many arguments against MENTALITY are inconclusive. This defensive strategy involves rejecting widespread internalistic assumptions about the nature of mental states; it draws on interesting parallels with philosophical arguments for wide content (Chapter 2, §1-3); it involves a well-known argument against luminosity of mental conditions (Chapter 4), along with an argument to the effect that mental states are, like knowledge, *prime* — i.e., non-decomposable into internal and external conditions (Chapter 3). Finally, it involves an argument against the claim that knowledge is not causally efficacious in the way mental state are (Chapter 2, §4).

While this whole package offers an outstanding defensive case for MENTALITY, the overall persuasiveness of the case is predicated on the presumption that MENTALITY ought to be the default view, to be dismissed only in the face of strong theoretical considerations to the contrary. But not many have granted that MENTALITY ought to be the default view. The general reaction

to Williamson's strategy is that the theoretical costs of the view far outweigh the initial apparent advantages. Some implications of the view that the mind is world-encompassing have appeared intolerable to many — especially its metaphysical consequences for the localization of the mind. If the truth of what is known by one supervenes on one's mental state, then a distal change in the environment can affect the mental state in a subject even if nothing in one's internal consciousness or neurology changes at that moment. For example, to use (Nagel, 2011, Section 2)'s often cited example, consider an ordinary well-informed American citizen who knows that Abraham Lincoln is President early in April of 1865 but whose mental state switches from knowledge to mere belief at some point on the evening of the 15th, as soon as the President is assassinated and the proposition that he is the President thereby becomes false. This person's mental state would have undergone this change, even if she had not heard what had happened in the faraway theatre, and even if she would have said and done just the same things immediately following this change as before it. Some find it unbelievable that one's mental state can change in this fashion. Relatedly, MENTALITY conflicts with the views that causal powers are located when their effects are located and that there cannot be action at distance, which many find compelling (e.g., Sosa (2009b), Jackson (2002), Fricker (2009), Smith (2017), Molyneux (2007), Magnus & Cohen (2003)).

I don't mention these worries because I share them or to cast doubt on MENTALITY so much as to indicate that there is a need, and room, for a more positive case for MENTALITY. A more positive case for MENTALITY would not take it as an initial presumption that knowledge is a mental state — as a claim to be given up only in light of compelling arguments against it. Rather, a positive case would emphasize the theoretical advantages of thinking of knowledge as mental. By doing so, one might motivate thinking of the theoretical costs listed above as consequences that we must accept, in a general cost-benefit analysis, if they are outweighed by the most substantial theoretical benefits.

Nagel (2011) has provided one such more positive case and I will return to considering it. But before, I would like to discuss Williamson's argument for the causal efficacy of knowledge — (Williamson, 2000, pp. 61-63) (cf. also Williamson (1997)) — for it might seem to provide a more positive direct case for MENTALITY. Here, Williamson purports to show that knowledge can be no less causally efficacious than other mental states, thus dispelling one more reason to dismiss knowledge as non-mental. While I think that this argument does succeed at showing that knowledge enters essentially in causal explanations of behavior — which is the only conclusion Williamson is aiming at in that chapter — I find this argument wanting if taken as a direct positive

argument for MENTALITY but for reasons quite different from those pointed out in the ensuing debate. This discussion will help me motivate and structure my argument ahead.

Here, Williamson takes issue with those who do not consider knowledge to be a mental/psychological state on the ground that knowledge does not ‘need to be cited in psychological explanations of action’ (cf. (Noonan, 1993, 291-2)). Thus, Williamson purports to show otherwise on the grounds that ‘suspicion is legitimate of a purported mental state, reference to which never plays an essential role in causal explanation’ (Williamson, 2000, Chapter 2, §4, p. 61). In effect, Williamson argues for EXPLANATORINESS:

EXPLANATORINESS Knowledge enters essentially in explanations of behavior.

The most famous example that Williamson offers on behalf of EXPLANATORINESS is the *burglar’s example*. A burglar spends all night ransacking a house, risking discovery by staying for such a long time. What explains his staying so long? Williamson suggests (p.62) that ‘a reasonable answer is that he knew that there was a diamond in the house.’ Granted that it is appropriate to invoke knowledge in an explanation of the burglar’s extended action, why think that knowledge enters *essentially* in this explanation?

Williamson’s response is that, in general, *the more closely connected the explanans and explanandum are, the better the explanation*. And the more closely connected the explanans and explanandum are, the *higher* the probability is of the explanandum on the explanans. In this particular case, Williamson observes that substituting ‘believe truly’ for ‘know’ would weaken the explanation, by lowering the probability of the explanandum conditional on the explanans: the probability of his staying all night ransacking the house conditional on his only having a belief, or a true belief, or a true and justified belief that is not knowledge, is lower than the probability of his staying all night conditional on his knowing that there are diamonds in the house. Why so? The answer, according to Williamson, is that knowledge is more stable (p. 63) than, e.g., *mere believing*, or than *merely believing truly*, and even than *believing truly and justifiably*. One respect on which knowledge differs from those weaker conditions is that it cannot be based on false premises. And its not being based on false premises makes it more resilient to possible defeaters than any weaker condition. Suppose the burglar was told that George possesses a diamond, and hence that there is a diamond in George’s house; however, it turns out that actually George’s wife possesses the diamond, so there is a diamond in George’s house, but not because George possesses it. After looking through George’s belongings and not finding the diamond, the burglar is likely to

stop ransacking the house. However, that would not happen if the burglar *knew* that there was a diamond in the house — in this case, evidence against there being diamonds in the house would be less likely to defeat the burglar’s enterprise. So, the probability of the explanandum being true on the explanans being true is higher if the explanans features knowledge than if it features a weaker condition than knowledge. If this measure is taken for explanatory loss, then any replacement of knowledge with a weaker condition would result in a worse explanation.¹ In this sense, knowledge appears essentially in explanation of the burglar’s resolute behavior, for replacing knowledge in it with weaker conditions would result in explanatory losses.

A first observation about this argument is that if successful, it would only vindicate a quite *limited* explanatory role for knowledge. The ‘added value’ of knowledge — its stability, i.e., the fact that it is harder to defeat than any weaker condition — is really only going to make a difference in explanation of some *extended actions*. Specifically, it will make a difference only in explanations of a particular class of extended actions — *resolute actions*, those manifesting the stability of knowledge in the face of possible defeaters. So, what the argument shows, at most, is that knowledge enters essentially in explanation of *resolute actions*. It is hard to shake the feeling that if knowledge were indeed a genuine mental state, its explanatory role would not be so limited.

The second observation is on the gap between EXPLANATORINESS, on one hand, and MENTALITY, on the other hand — from the conclusion that knowledge enters essentially in some explanations of behavior to the conclusion that knowledge is a mental state. While it is reasonable to expect that mental states (or mental properties) enter in causal explanations of behavior, it is not true that if a state (or a property) enters in a causal explanation of behavior, it is then a mental state (or a mental property). For example, consider explaining Mark’s habit of *speaking loudly*. Mark’s speaking loudly might call for a purely psychological explanation — e.g., in terms of his confidence. However, in some explanatory contexts, a *hybrid* explanation — both psychological and not psychological — might do better: it might be that Mark’s speaking loudly is to be explained both in psychological terms, in terms of his confidence, and in non-psychological terms, in terms of the status of his vocal cords. Indeed, the probability of his speaking loudly is higher on the hypothesis that he is both very confident and his vocal cords allow for high volume than just on the hypothesis that he is very confident. Why think an explanation of the burglar’s extended action is not hybrid in the same way — involving both a psychological condition (belief) and a non-mental property of that belief (it amounting to knowledge)?

¹Though see Molyneux (2007) for a criticism of Williamson’s proposed criterion for good explanations.

Indeed, many might concede that epistemic states, such as being justified or warranted, can figure essentially in certain explanations, without having to conclude that they correspond genuinely to mental states. For example, my belief that Mark is guilty of murder being well supported (let us assume, an epistemic property of my beliefs) might explain my success at convicting Mark and might explain it better than if I just had the same belief but that belief was unsupported. However, an explanation that essentially appeals to the epistemic property of being supported does not necessarily show that being supported is a mental state (or a mental property) in its own right, over and above having the belief that I have. Unless the relevant explanation is one that we have independent reasons to think belongs — or ought to belong — to psychology, knowledge’s appearing in explanations of certain features of behavior is compatible with it being non-mental — or with it being mental only derivatively — for non-mental properties and states can enter essentially in explanations of behavior.

Thus, even if Williamson’s burglar argument succeeds in its own terms — in that it succeeds at establishing EXPLANATORINESS — it falls short in providing a direct argument for MENTALITY. What would a positive argument for MENTALITY look like?

2.2 Nagel on Mentality

Nagel (2011) (cf. also Nagel (forthcoming), Nagel (2017)) has made a formidable case for the claim that as matter of fact, knowledge does enter in explanations that psychologists themselves deem as falling within their domain. She points out that developmental, social, cognitive, and comparative psychologists explicitly and without hesitation classify knowledge as a mental state, alongside states of belief, desire, intention, perception and feeling (e.g., Baron-Cohen et al. (1994); Call & Tomasello (2011); De Villiers (2007), Epley & Waytz (2009); Heyes (1998); Keysar (2007), Lin et al. (2010); Premack & Woodruff (1978), Saxe (2005); Sodian et al. (2006), Wellman & Liu (2004)). If so, then a knowledge-less psychology would be highly revisionary of ordinary psychological practice: giving up the mentality of knowledge would commit one to an error theory of wide branches of psychology. Her work has also provided ample evidence that the role of knowledge in explanations of behavior that psychologists themselves offer goes well beyond an explanation of extended actions, therefore improving on the question of scope remarked above.

One might, of course, worry that psychologists actually appealing to knowledge in explanations of behavior is not sufficiently principled — i.e., that while they do, as a matter of fact, appeal

to knowledge all the time, this use of knowledge is not essential, and those explanations could be rephrased without loss by replacing knowledge with some weaker condition. However, Nagel's argument for MENTALITY is not just a matter of registering a not well-known fact about psychologists' practices. Both Nagel (forthcoming) and Nagel (2017) discuss independent empirical evidence in favor of the priority of the concept of knowledge over that of belief, such as evidence from developmental psychology that the former is learned before the latter (Bartsch & Wellman (1995), Tardif & Wellman (2000), Shatz et al. (1983)); evidence from both human and nonhuman animal psychology that suggests that knowledge is easier to be recognized than both false beliefs and true beliefs that fall short of knowledge, and that our interpretation of others' behaviors is influenced by default expectations that they act knowingly, rather than otherwise (Papafragou et al. (2007)); that the false belief task can be read as showing a tendency to over ascribe knowledge over weaker conditions (Hogrefe et al. (1986)); she, further, overviews recent work on nonhuman primates indicating that when interpreting behavior, they can come to understand what others see, hear and know (Call & Tomasello (2011)), and presents reasons for preferring this interpretation of the findings over others that play down the mind-reading abilities of those primates. Finally, Nagel (2017) has given several reasons to think that the concept of knowledge is a helpful tool in our mind-reading practices, which go a long way to show the importance of knowledge ascriptions in a theory of mind-reading.

These are powerful abductive considerations in favor of MENTALITY, that considerably buttress its plausibility. Yet, a different sort of concern looms. As (Nagel, 2011, Section 4) is well aware, even if our natural mind-reading abilities avail themselves to the epistemic mental state *concepts* of knowledge and belief, rather than of that of belief alone, mental reality might still fail to correspond to our natural way of tracking it. In order to display this point more vividly, compare the case of knowledge to that of causation. An *eliminativist Humean* might think that a description of the physical world would not have to involve the relation of causation and would still grant that the *concept* of causation, as well as the habit of noticing correlations, can still be very helpful as heuristics to latch onto correlations that are important for our purposes to single out. A Humean of this brand might allow that the concept of causation figures prominently in our explanations of the correlations that we can observe, without granting that causation must be therefore (fundamentally) real.² Nagel (2011)'s argument fails to rule out a corresponding form of *knowledge eliminativism*, on which the concept of knowledge plays an important role in our mind-reading practices and so

²For discussions of Humeanism, see Lewis (1974), Loewer (2012), Frisch (2020).

in the prediction of the behavior of others but which denies that knowledge is mentally real.

Because of this, while Nagel has made an outstanding positive case for MENTALITY, I believe that the case for MENTALITY can be further strengthened. I believe that there is a positive case to be made for MENTALITY that centers around the importance of knowledge in a theory of *agentive control* — a property of performances that psychologists and philosophers of cognitive science deem as central to skillful behavior. I will return to assessing the advantages of this argumentative strategy in the conclusion. Here, I can anticipate that, so framed, my discussion will enable me to make a plausible case for MENTALITY, which will be able to bypass issues regarding what makes an explanation of behavior genuinely psychological that arise for Williamson (2000)'s strategy, as well as worries concerning the plausibility of inferring the mentality of knowledge from the usefulness of the concept of knowledge in our mind-reading practices that arise for Nagel (2011)'s strategy.

3 Agentive Control in the Psychology of Skills and in Action Theory

At least since the Seventies, the notion of 'control' has been invoked in psychology and cognitive sciences (Shiffrin & Schneider (1977), Hommel (2000), Posner et al. (2004), Palmeri (2006)) and used to refer to a general property of cognitive and motor processes. More recently, cognitive scientists as well as philosophers of cognitive science and psychology have been theorizing in particular about so-called *agentive control* — which is taken to be a distinctive property of *skillful behavior*.

So, for example, Miller (2000) claims that 'cognitive control is essential for what we recognize as intelligent behavior' and that 'agents have evolved mechanisms that can override or augment reflexive and habitual reactions in order to orchestrate behavior in accord with our intentions, which exert cognitive control over lower-level sensory, memory, and/or motor operations' (Miller, 2000, p. 59). According to Miller (2000), this cognitive control is characteristic of highly intelligent performances, such as those exhibited by experts' behaviors in both motor and cognitive tasks. Dreyfus (2002) thinks that control is distinctive of expert performance and tells us that an expert is in control of their movements in the sense that they can stop doing what they are doing if they so want. Christensen et al. (2016) argue that cognitive control is present in expert skillful

performance in the form of ‘higher-order action control’. Wu (2016) focuses on *agentive control*, which he thinks ‘yields phenomena of central philosophical interest: moral, rational, reason-based, skilled, conscious, epistemic and free agency’ (Wu, 2016, p. 101). Verbruggen & McLaren (2014) argue that adaptive and goal-directed behavior is to be understood in terms of agentive or executive control, in virtue of which people can regulate their behavior according to higher-order goals or plans. Finally, Toner et al. (2015) argue that skilled athletes use cognitive (or executive) control to maintain or improve their performance proficiency (see also Christensen et al. (2015), Fridland (2014), Christensen et al. (2019), Shepherd (2014), Wu (2011)). These psychologists and philosophers all seem to consider as a hallmark of skillful performance that it is, in some sense, to be specified, under the agent’s control. They all seem to endorse:

CONTROL CONSTRAINT ON SKILLFUL ACTION Whenever an agent ϕ s skillfully,
they are in control of their ϕ -ing.

According to CONTROL CONSTRAINT ON SKILLFUL ACTION, skillful performance is *controlled* performance. Following common usage, I will call this property in virtue of which skillful performances are controlled *agentive control*. Agentive control is *factive* in the sense that if an action of an agent possesses it, then that action is successful — failed attempts cannot be controlled.

The notion of agentive control does not just play a central role in emerging accounts of skillful action. In terms of philosophical literature, it actually makes its first entrance in the philosophical debate on the nature of *intentional action*. Many think it is a platitude about intentional action that it is — in some sense — under the agent’s control. As an illustrative example, Ryle (1949) contrasts the clown, who tumbles intentionally for an audience, with a klutz, who trips and tumbles inadvertently (Ryle, 1949, p. 33). It seems that their actions are distinguished — in part — by the fact that the clown is in control of their tumbling, whereas the klutz is not. This motivates:

CONTROL CONSTRAINT ON INTENTIONAL ACTION Whenever an agent ϕ s intentionally,
they are in control of their ϕ -ing.

Further motivation for the CONTROL CONSTRAINT ON INTENTIONAL ACTION comes from its ability to explain a wide variety of cases that have loomed large in the philosophy of action literature. In this literature, control conditions on intentional action have been proposed in order to impose anti-luck, or anti-accidentality, conditions (Gibbons (2001)) and to overcome deviant causal chains (Schlosser (2007), Schlosser (2008)). Consider cases involving deviant causal chains. It is a

familiar observation that an agent can intend to ϕ , and that this intention can lead them to ϕ , but that they nonetheless do not ϕ intentionally, because the intention and its execution are not connected in the “right way” (Davidson, 1980, p.79). A standard example (from Chisholm (1966)):

KILLING NEPHEW A nephew intends to murder his uncle, but this unnerves him so much that he drives excessively fast and accidentally kills a pedestrian who turns out to be his uncle.

The CONTROL CONSTRAINT ON INTENTIONAL ACTION explains this intuition. While the nephew’s intention does cause the fast driving and the accident, his action is clearly not under his control.

Agentive control is plausibly a common feature of both skillful and intentional action. Beddor & Pavese (2020) argue that this is not surprising: after all, skillful actions are paradigmatically intentional. Consider Ryle (1949)’s example once again: while the performance of the clown, who tumbles intentionally for an audience, is skillful, that of a klutz, who trips and tumbles inadvertently, is not skillful (Ryle, 1949, p. 33). If agentive control is distinctive of skillful and intentional action, any theory of skillful and intentional action will have to cast light on agentive control. Beddor & Pavese (2020) go on to present a theory of skill that is capable of explaining this striking commonality between skillful action and intentional action.

The focus of this paper is not a theory of skill but rather a theory of agentive control and its consequences for MENTALITY.³ So let us ask: what does it mean to have control over one’s action — in virtue of what an action is controlled? Some explicitly refrain from providing a theory of agentive control and take it as a primitive feature of performances or cognitive structures (e.g, Fridland (2021)). However, I submit that it would be preferable not to simply import notions from psychology as primitives. We want to understand what those psychological notions amount to, if they are then put to do work in our philosophical theories of intentional action and skillful action. In the next section, I consider a notable attempt at understanding agentive control in terms of psychological mechanisms, such as attention and focus. I argue that attempts of this sort are doomed to fail, but that it is instructive to understand why they do so.

³Because it is not essential that control is necessary for intentional action, this paper will not proceed on the assumption that every intentional action is controlled. What is important for my purposes is that there is such a thing as agentive control, and that this property characterizes at least skillful performances and maybe also intentional performances.

4 The Limits of Psychological Theories of Control

It is natural to expect that a theory of agentic control should invoke psychological processes that agents undergo when they perform skillful and controlled actions. And many cases of agentic control seem to involve the psychological mechanism of *attention*. While the importance of attention for controlled action is acknowledged by many (e.g. Shiffrin & Schneider (1977), Hommel (2000), Posner et al. (2004), Palmeri (2006), Montero (2010), Montero (2016), Toner et al. (2015)), Wu (2016) has offered the most systematic theory of agentic control in terms of the psychological mechanism of ‘intention-attention coupling’ (cf. also Wu (2011)).

According to Wu (2016), central to a theory of agentic control is to provide an account of deviant causal chains, illustrated above by **Killing Nephew**. In deviant causal chains, the success, though intended, is accidental, so it is neither intentional nor skillful. Therefore, agentic control seems to be missing in deviant causal chains but is present in standard cases of intentional action. If this holds true, then a theory of agentic control needs to be able to overcome the problem of deviant causal chains, and by doing so cast light on what it takes for an action to be intentional.

Wu (2016)’s innovative proposal is to define agentic control in terms of a nexus between intention and attention: ‘the problem with current causal accounts is that they ignore a basic psychological capacity essential to control: attention’ (Wu, 2016, p. 102). If attention is critical to agentic control, then these theories are at best incomplete and at worst inaccurate.’ According to Wu, integrating attention into the causal theory provides an analysis of the basic case of agentic control that shows what goes wrong in classic examples of causal deviance and illuminates specific forms of agency, such as skill and expertise. In typical cases of intentional action, where an agent is performing a bodily action in relation to some object in the environment, the subject’s experience of the object informs a response. This ‘coupling’ of the experience and the response involves selecting an action: specifically, the subject selects the object for the action — for example, reaches out to a cup in their visual field. Where we have a specific aspect of the subject’s experience operating in this way, we have a selection of what is experienced for action. This coupling of the subject’s experience and the selection is a form of attention.

The idea that Wu (2016) pursues then is that if the intention-attention nexus is crucial to agentic control, then defects in that nexus can illuminate defects of agency as in causal deviance, and indeed other defects as well (e.g. distraction). How does this theory of control overcome the problem of deviant causal chains? Let’s return to **Killing Nephew**. What goes wrong in this sort of

example, according to Wu, is that the nephew's intention does not lead to appropriate 'coupling' with the result that his movement fails to be guided by attention to the target. The resulting killing is accidental. For there to be an intentional killing, the paradigm case would be that the nephew intends to drive over his uncle, namely that pedestrian, requiring coupling of attention to the spatial location of his uncle to inform his driving response. But appropriate coupling is clearly missing despite the causal influence of intention on the process. According to Wu, this is an absence of agentive control because of a disruption in the appropriate intention-attention nexus. So by highlighting the intention-attention nexus, we illuminate agentive control. On that basis, Wu explains how many classical cases of deviance amount to failures of control.

While Wu is certainly right that in many paradigmatic cases of intentional and skillful action, psychological mechanisms, such as attention and its nexus with intention, play a central role, there are cases of controlled action in which attention and perception — and more general monitoring mechanisms — play absolutely no role at all. I will mention two classes of cases as examples. First, consider **intentional omissions**. We often intentionally do not do things. For example, today I intentionally did not shovel the patio, despite it being fully covered in snow. I intended not to do it (I had even deliberated about it) because I had other things to do (such as writing this paper) that I deemed more urgent. In intentional omissions of this sort, attention plays little if any role. I intentionally did not shovel the patio the entire afternoon, but I spent almost no time this afternoon attending to my patio or shoveling. The second class of cases are **deferred intentional actions**. A mob boss might intentionally kill a member of the opposite clan by deferring to others to complete the murder. If the mob boss chooses his helper wisely, the mob boss is definitely in control of the killing. But that is true whether or not the mob boss pays much attention to the performance of the deferred act, beyond selecting the right helper.

Notice that the problem of deviant causal chains can arise for omissions and deferred actions too. Suppose that although I intended not to shovel the patio, that somebody would have forced me to do it, were it not for a fortuitous sequence of events (e.g., my neighbor tried to force me out of the house to shovel, threatening me with a gun, but he slipped on ice while doing so and hit his head, causing him to die). In this case, I would not be in control of the execution of my intention, but not because I did not pay attention to its execution. Perhaps more clearly, deviant causal chains can be reproduced for deferred actions. Suppose the mob boss selects as his hitman somebody that has no intention to follow his orders but who, for accidental circumstances, ends up killing the right person. Intuitively, the mob is not in control of the killing and the killing was not intentional

(neither the deferred killing nor the actual killing) on his part. And yet, it is not because he did not pay enough attention — he would not have paid attention to the killing even if the helper had turned out to be loyal.

The conclusion is, I think, that controlled action is more general than attention-guided action. So, a theory of agentic control in terms of the intention-attention nexus is unlikely to solve the problem of deviant causal chains in its generality. While the causal theory of intentional action is indeed afflicted by deviant causal chains, the solution is not to integrate it with a theory of the role of attention in action. A solution must be more general than that.

5 Controlled Action First?

What should a theory of agentic control look like? In this section, I discuss some reasons for being skeptical of the project of reducing controlled action to simpler conditions. In the following section, I argue that controlled action is prime and conclude that it is analyzable only in terms of other prime conditions. That will set the stage for arguing that only an epistemic theory of agentic control provides a plausible generalization of the psychological theories of agentic control discussed in the last section.

Several philosophers have recently warned against attempts to understand intentional action in terms of more basic notions. If their concerns were sound, that would underscore some skepticism towards the project of putting forward an epistemic theory of agentic control outlined, for a theory of agentic control naturally yields a theory of intentional action. According to (Levy, 2013, p. 705):

The reductive program in the philosophy of action is unlikely to bear fruit, and should be expunged from the philosophical agenda. Intentional action is not susceptible to reductive decomposition; it should be treated as basic.

This same point is made by (Williamson, 2017, Section 3), in a paper that investigates the relation between knowledge and action:

The attempt to solve the equation for action generated an extensive but still unsuccessful research program. Many solutions were proposed, of increasing complexity, but unless circular always eventually succumbed to counterexamples, of increasing

complexity, to their necessity or sufficiency. The world-to-mind direction of fit for intention suggests an instructive case in point: causal theories of action. At its simplest, such a theory identifies Y with the condition that $S \phi$ s because (causally) S intends to ϕ . Such theories succumbed to the problem of deviant causal chains: the causation from intention to success may be of the wrong kind for action. For example, my intention to drop the Ming vase causes my hands to sweat and become slippery, which in turn causes me to drop the vase. The three conditions (A1)-(A3) hold on the proposed reading, but I do not intentionally drop the vase. Attempts are still made to solve the problems for attempted analyses of action by adding epicycles, but none have worked.

Finally, in discussing the prospect of analyzing intentional action in terms of knowledge, Moss (2020) echoes almost exactly (Williamson, 2017, Section 3).

As far as I can see, there are two different arguments that underlie skepticism for the prospect of a reductive analysis of intentional action. The first argument relies on the analogy with Williamson's argument for the un-analyzability of knowledge. Here, the idea is that Williamson's reasons for thinking that knowledge cannot be analyzed can be reproduced exactly in the case of intentional action. The second argument is an inductive skepticism argument: because no reduction has succeeded so far, then no reduction should be expected to succeed in the future.

I don't have much to offer by way of a response to the inductive skepticism argument, beyond proposing in due course a theory of agentive control that strikes me as promising. But I do want to address the first point — the one according to which Williamson's main argument for the un-analyzability of knowledge can be reproduced to argue for the un-analyzability of intentional action.

Williamson (2000)'s argument for thinking that knowledge cannot be analyzed can be reconstructed as consisting of two distinguishable parts. The first part is the argument that knowledge cannot be decomposed into external and internal conditions — the argument that knowledge is “prime” (Williamson, 2000, pp. 65-93). The second part is the argument that, among the prime conditions, knowledge is the most general one — the argument to the effect that knowledge is the most general factive state (Williamson, 2000, pp.33-41). Both parts are needed for Williamson's argument to be successful: the first part only establishes that knowledge cannot be analyzed in the sense that it cannot be decomposed into internal and external components. But even if knowledge were not decomposable into internal and external components, the prospect of analyzing knowledge in terms of other more basic conditions would still be open, since it would still be open to

analyze knowledge in terms of other *prime* states.⁴ This sort of analysis, if feasible, would not involve a decomposition of knowledge into external and internal conditions. So, in order to reach the conclusion that knowledge is not analyzable *tout court*, the second part of the argument is also needed — an argument to the effect that knowledge is the most general factive mental state.

Both parts of the argument have been challenged, of course: virtue epistemologists have objected that not every analysis of knowledge needs to take the form of a decomposition into external and internal components — and indeed virtue epistemological analyses of knowledge are not analyses of this sort (e.g., Sosa (2009a), Sosa (2007), Beddor & Pavese (2018)). Others have challenged the claim that knowledge is the most general factive mental state. It is not my goal to assess or even defend Williamson’s argument for un-analyzability of knowledge. But I’d like to point out that attempts to run a parallel argument for the un-analyzability of intentional and controlled action are doomed to fail. That is so because a parallel argument for the un-analyzability of intentional and controlled action would *also* have to consist of two distinguishable parts: the first part would be to establish that intentional and controlled action is ‘prime’ — that is, that it cannot be decomposed in terms of an internal component (intention, beliefs, knowledge-how etc.) and an external component (whatever external condition allows for the action’s success). The second part would be to argue that there is no other prime condition that controlled action can be reduced to. In the next section, I will argue that controlled action is indeed prime. However, this would not establish that intentional action is not analyzable *tout court* — only that it cannot be decomposed into a conjunction of an internal and an external condition. In order to reach the conclusion that intentional and controlled action cannot be analyzed, one would need, in addition, the second part of the argument — showing that intentional action is the most general prime condition, and that there is no other prime condition that intentional action can be analyzed in terms of. But controlled action and knowledge cannot *both* be the most general prime condition. And among the two, controlled action is more likely to be in need of reduction than knowledge is, if not for anything but the fact that only mental states have their intentionality intrinsically. So, even if both knowledge and controlled action are prime, that by itself provides little ground for thinking that controlled action cannot be analyzed, especially in terms of knowledge.

⁴Indeed, Sosa (2009b) has challenged the latter step of the argument, claiming that knowledge might be understood in terms of safety, which is also a prime condition.

6 Primeness of Agentive Control

In this section, I argue that controlled action is prime — in Williamson (2000)'s sense that it cannot be decomposed into internal and external conditions. If that is correct, this supports thinking that if analyzable at all, agentive control can only be analyzed in terms of other prime conditions.

Consider the idea that controlled actions are decomposable into internal and external conditions. What would those conditions be? Presumably, the intention or the attempt that gives rise to the action feature among the internal conditions, whereas it would be the bodily movements and the environmental conditions that enable the success feature among the external conditions. However, there are reasons to think that controlled actions are not easily decomposable into internal and external conditions in this manner. Here is a first, rather simple-minded, argument for this conclusion. If attempts exhausted the mentality of actions, then provided that one attempted to perform an action ϕ , one's eventual success at ϕ would have to be intentional. For on this picture, the intentionality, and hence the mentality, of actions would be exhausted by their attempts. However, there are a variety of cases in which one attempts at ϕ -ing, succeeds and yet their action fails to be intentional or controlled. Lucky successes are examples of this. Consider Mele & Moser (1994)'s **Nuclear Reactor**'s case. A nuclear reactor is in danger of exploding. Fred knows that its exploding can be prevented only by shutting it down, and that it can be shut down only by punching a certain ten digit code into a certain computer. Fred is alone in the control room. Although he knows which computer to use, he has no idea what the code is. Fred needs to think fast. He decides that it would be better to type in ten digits than to do nothing. Vividly aware that the odds against typing in the correct code are astronomical, Fred decides to give it a try. He punches in the first ten digits that come into his head, in that order, believing that his doing so might thereby shut down the reactor and prevent the explosion. What luck! He punched in the correct code, thereby preventing a nuclear explosion. This action is not under Fred's control, despite being guided by his intention to defuse the nuclear reactor. So, the mere fact that one can have non-controlled actions that are guided by one's intentions suggests that controlled action is not easily decomposable to internal and external conditions.

In order to show in a more systematic fashion that controlled action is a prime condition, we can import the set-up from Williamson (2000), who gave us a general recipe for showing that a condition is prime (Williamson, 2000, Chapter 3, pp. 65-92). Some definitions first. Let a case be a centered world. And let a condition be something that can obtain or not in a case. A condition can

be mental or non-mental (physical). Conditions (mental or physical) might be narrow or broad. Say that a case α is internally like a case β if and only if the total internal physical state of the agent in α is exactly the same as the total internal physical state of the agent in β . A condition C is narrow if and only if for all cases α and β , if α is internally like β , then C obtains in α if and only if C obtains in β . C is broad if and only if it is not narrow. A case α is externally like a case β if and only if the total physical state of the external environment in α is exactly the same as the total physical state of the external environment in β . A condition C is environmental if and only if for all cases α and β , if α is externally like β , then C obtains in α if and only if C obtains in β . C is prime if and only if it is not composite. A narrow mental condition is trivially composite, for it is the conjunction of itself with the environmental condition that holds in all cases whatsoever. A condition C is composite if and only if it is the conjunction of some narrow condition N with some environmental condition E. C is prime if and only if it is not composite.

Case	Way	Internal	External	Joint	Result
α	1	✓	✓	✓	✓
	2	✗	✗	✗	
β	1	✗	✗	✗	✓
	2	✓	✓	✓	
γ	1	✓	✗	✗	✗
	2	✗	✓	✗	

Figure 1: Recipe for

Primeness

Given this set-up, how can we show that a certain condition C is prime? Here is a general recipe that (Williamson, 2000, p. 68) has taught us (cf. Figure 1). We can imagine circumstances in which C can be realized in just two ways, which need not be mutually exclusive. One is in C if and only if one is in C in either way 1 or way 2. Each way involves a channel with an internal and an external part; one is in C in way i if and only if both the internal and the external parts of way i are open. In case α , both the internal and the external parts of way 1 are open but neither the internal nor the external part of way 2 is open; thus, one is in C in way 1 although not in way 2; therefore, one is in C. Case β reverses the two ways in status. In β , neither the internal nor the

external part of way 1 is open but both the internal and the external parts of way 2 are open; thus, one is in C in way 2 although not in way 1; therefore, one is in C. In case γ , the internal part of way 1 but not the internal part of way 2 is open, because γ is internally like α , and the external part of way 2 but not the external part of way 1 is open, because γ is externally like β ; thus, one is in C in neither way 1 (because its external part is not open) nor in way 2 (because its internal part is not open); therefore, one is not in C. The relations between α , β , and γ ensure that the condition that one is in C is prime.

With this recipe, it is easy to show that controlled actions are prime. Consider, for example, Ben's skillful action of jumping to a new height, which Ben can realize in two ways: either by employing Ben's standard jumping technique but by exploiting a particularly favorable wind; or by a new technique that requires his best physical condition but not necessarily a favorable wind condition. Let α be a condition in which internally Ben intends to jump to a new height by exploiting the favorable wind and knows exactly how to do so; externally, the wind is indeed favorable. Ben skillfully jumps to a new height in α . Let β be a condition in which internally Ben plans and attempts to jump to the new height with a new strategy that requires for his best physical condition; externally the wind is not particularly favorable but the new strategy does not require particularly favorable wind conditions and, further, Ben is in good physical shape. Ben skillfully jumps to a new height in β . Finally, take γ to be internally like α in that Ben intends to jump to a new height by exploiting the favorable conditions wind and knows how to do so; externally, it is like β in that Ben is in good physical condition, but the wind is not favorable. In γ , Ben does not skillfully jump to a new height (maybe he simply fails to deliver the result; or he succeeds but not skillfully and only due to lucky circumstances, as the unfavorable wind turns out to help his performance).

Notice that nothing hinges on the details of the case — for pretty much every controlled action, however we specify the relevant internal conditions, we can envision two ways of performing it and three cases α , β , and γ such that γ agrees internally with α and externally with β but where the action is not controlled. If this is correct, then just like knowledge and other broad conditions, controlled action is prime — non-decomposable into internal and external conditions.

7 An Epistemic Theory of Agentive Control

If controlled action is prime, then we cannot hope for a theory of agentive control that does not appeal to other conditions that are also prime. In this section, I argue that an epistemic theory of con-

trol that combines a practical knowledge condition with a knowledge-how condition is promising and provides a desirable generalization of the psychological theories of agentic control discussed earlier.

The psychological theories of agentic control I discuss try to understand agentic control and non-deviancy in terms of monitoring capacities, such as attention. As we have seen, the attention-intention nexus is unlikely to enter in an explanation of every case of agentic control. For example controlled omissions and controlled deferred actions are not attention-guided. More generally, the underlying idea that *agency* ought to require monitoring is dubious. Even if many cases of controlled action involve monitoring them, it seems that we can at least imagine agents who can be in control of their actions even though they don't constantly monitor them (perhaps deferred actions are the only actions they are capable of). If so, monitoring and the intention-attention nexus are at best only contingently associated with controlling agency.

So, these psychological theories of agentic control are too limited in scope. However, their considerations suggest a plausible generalization. Attending to our performances are *ways in which we can come to know what we are doing while we are doing it* —i.e., they are capacities for 'practical knowledge' in a broadly Anscombean sense.⁵ So, it might be that what is distinctive of controlled action is not so much attention-guiding but rather the practical knowledge that attending to our actions give rise to. While in many cases of intentional performance, we come to know what we are doing while we are doing it because we are guided by attention and other monitoring mechanisms, in other cases of intentional performance, we come to know what we are doing while we are doing it *through other means*. For example, consider again **Deferred Actions**: although the mob boss does not need to attend to the killing of his enemy when it happens for that action to be under his control, the mob boss might still know that he is killing his enemy through his hitman, provided that he knows that his hitman is loyal and will not fail. Similarly, consider the case of **Intentional Omissions**: I don't have to attend to my omitted shoveling to know that I am not shoveling snow off my patio. Though I don't need to monitor my not shoveling the patio right now, I can still know that I am not shoveling the patio for the entire length of the afternoon

⁵While my views are deeply shaped by Anscombe (1963)'s view of practical knowledge, I do not intend my use of practical knowledge in an account of agentic control to be faithful to Anscombe's understanding of practical knowledge. For one thing, for Anscombe, practical knowledge is non-observational. By contrast, the sort of knowledge monitoring mechanisms can give rise to is plausibly observational. Moreover, for Anscombe, practical knowledge causes the action to come into existence; by contrast, my account is neutral on the explanatory relation between practical knowledge and the corresponding action.

when I omit to shovel the patio. So, this suggests a plausible generalization: the relevant notion of agentive control can be cashed out in epistemic terms.

As Anscombe (1963) and others have pointed out, this seems quite intuitive: it sounds extremely odd to say, ‘She has things under her control but she has no idea what she is doing’ (Anscombe, 1963, p. 11). Indeed, it seems a platitude that being in control of one’s action is, at least in part, a matter of knowing what one is doing when one is doing it. An indication of this is that we can, and often do, express control over our endeavors by saying that we know what we are doing. This suggests that practical knowledge is necessary for agentive control:

Epistemic Condition on Agentive Control (ETC) Someone is in control of ϕ -ing at time t only if

Practical Knowledge they know what they are doing when ϕ -ing at t .

This necessary condition is certainly not sufficient. One might come to know what one is doing at a time without doing it intentionally at that time. Consider the clumsy person who tumbles unintentionally. While tumbling, they might come to know that they are tumbling (perhaps they are tumbling in front of a big mirror). And yet clearly they have no control over it. It would not be sufficient to add that one *intends* to perform the relevant action and that one is guided by that intention. Consider a case in which one intends to perform a ski stunt but has false beliefs about how to perform it. However, because of a neurological condition, which translates the intended movements into the correct ones for performing that ski stunt, one comes to perform the action successfully and knows that one is doing it while one does it. This performance is not under the person’s control either, even though the agent is guided in their performance by their intention and by their practical knowledge. It seems that in order for agentive control to obtain, in addition to knowing what one is doing, one ought to know this *in virtue of* exercising their intention to perform the action and *in virtue of* being guided by one’s knowledge-how. So, while there are different ways of fleshing out an epistemic theory of agentive control, an account along the following lines holds considerable promise:

Epistemic Theory of Agentive Control (ETC) Someone is in control of their action at a time t if and only if

(i) **Practical Knowledge** they know what they are doing at t , and

(ii) **Intention-Know-How** they know this at t in virtue of exercising their intention to perform the action and in virtue of being guided by their knowledge of how to perform the action in question.

This condition integrates two ideas about agentic control. The first is that controlled action requires practical knowledge of what one is doing (e.g., Anscombe (1963), Hampshire & Hart (1958), Thompson et al. (2011)). The second is that controlled action requires some form of know-how or skills (e.g., Mele & Moser (1994), Mele & Svedlik (1996), Small (2012), Gibbons (2001), Pavese (2018), (Piñeros Glasscock, 2019, pp. 1242)). ETC combines these two ideas, requiring that one knows what one is doing in virtue of one's exercising both their intention and their know-how.

Several clarifications are in order. First, I consider the phrase 'in virtue of' to denote a grounding relation — a *partial grounding* relation to be exact. If one knows what one is doing at any time, that knowledge is, at least in part, grounded in the fact that one is doing it. So practical knowledge cannot be fully grounded in terms of one's knowledge-how. Yet, according to ETC, practical knowledge of what one is doing is partially grounded in one's doing it and partially grounded in one's knowing how to do it and in one's intention to do it. Second, consider the practical knowledge requirement in (i). Notice that it only requires that one knows that one is doing ϕ when ϕ -ing intentionally, not that one knows that one is ϕ -ing intentionally. The stronger requirement has been shown to give rise to paradoxical conclusions, by running a variant of Williamson's anti-luminosity argument, on the assumption that the margins for error principle holds for knowledge (cf. Piñeros Glasscock (2019)). But as argued in Beddor & Pavese (2021), the weaker requirement does not lead to the same paradoxical conclusions.⁶

Next, what is the content of the relevant knowledge in (ii)? I have argued elsewhere that this is knowledge of the means for performing an action (Pavese (2013), Pavese (2015), Pavese (2017)), and I have motivated this claim by showing that it follows naturally from the idea that intentional action requires an action plan, specifying the means for performing it (cf. Goldman (1970), Pavese (2018), Pavese (2020), Pavese (2021)). The means can be partial, just like action plans can be partial (cf. Bratman 1987), and the relevant knowledge can be probabilistic in Moss (2018)'s sense

⁶As Beddor & Pavese (2021) argue, one could still reformulate a Williamson's style anti-luminosity argument against such a weaker requirement but only on the assumption that there are *essentially intentional actions*. However, Beddor & Pavese (2021) provide reasons for thinking that the claim that there are essentially intentional actions is independently quite implausible.

(cf. Pavese (2020)) — it requires not that the means guarantee the successful execution of the action but that they are sufficiently likely to bring it about.⁷

ETC correctly predicts that when an action is caused deviantly, the agent is not control of it. Consider **Killing Nephew**. Here the nephew does not even know that he is killing his uncle, for he does not know that the pedestrian is his uncle. So, the practical knowledge requirement is not satisfied. Moreover, *even if* he did know that, he would not know it *in virtue of* exercising his intention to kill his uncle. Thus, ETC predicts correctly that the action in **Killing Nephew** is not intentional.

ETC seems to provide a plausible antidote to deviant causal chains. At this point, however, some might object that a less demanding theory of agentic control might work just as well. Why not say that control over one's action merely requires belief, or perhaps even justified true belief, about what one is doing, belief which is sustained by one's knowledge of how to perform the act in question and by one's intention? In order to answer this objection, it is helpful to consider both theoretical reasons for thinking that nothing weaker than knowledge will do and intuitive judgements about actions based on Gettiered beliefs.

First, the theoretical reasons for thinking that nothing weaker than knowledge will do. Agentic control can be given a modal specification: if one is in control of ϕ -ing, then in some sense to be specified, one could not easily fail at ϕ -ing. If this or a similar modal condition holds of controlled action, then we have grounds to think that any condition weaker than knowledge would not be adequate to capture agentic control. After all, justified true belief or true belief can be modally weak — they can easily be believed falsely (Sosa (1999), Williamson (2000)). So, if one bases one's action on a Gettiered belief about the means for performing, or if the performance is accompanied by a Gettiered practical belief about one's performing it, one could easily fail at the

⁷Is the intention requirement needed? Some have argued against intentional action requiring intention, on the basis of cases where side effects are unintended but intuitively count as intentional (cf. Ginet (1990), Bratman (1981), Knobe (2003)). Although a discussion of these cases would bring me too far afield, I actually don't think that these counterexamples to intentional action requiring intention succeed. When side effects count as intentional (such as wakening one's neighbor in the process of starting the car to go to work), it is because the side effects were reflectively considered and the agent concluded, through practical reasoning, that one ought to allow those side effects to happen if they want to reach a certain goal. This conclusion of practical reasoning is very much like an intention, though conditional. While I think that there is a real distinction between primary intentions — goals that are intended — and secondary intentions — side effects that are intended — I do think that every intentional action is intended either primarily or secondarily.

action. If agentic control entails modal robustness, then the sort of beliefs that make agentic control possible also ought to be subject to an anti-luck condition. This is a general theoretical consideration for thinking that knowledge can capture agentic control better than any weaker condition — that only knowledge has the right modal profile (for a similar argument, cf. also Greco (2016) and Pavese (2018)).

This general tenet can be strengthened by drawing on more intuitive considerations from cases where the agent acts on a Gettiered belief about what they are doing. Gibbons (2001) has provided one such case:

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 that if she buys a particular ticket that is handed to her, she will win. She buys the ticket and wins. So her belief that she will win the lottery by buying that particular ticket is true. It is even justified. Buying a winning ticket is a perfectly reliable way of winning a lottery and the lottery is not rigged but fair.

Here, Cindy intends to win the lottery and attempts to do so by buying a ticket that she believes truly and justifiably to be a winning ticket. And yet, Cindy's victory is neither controlled nor skillful nor intentional: the lottery turns out to be fair, and fair lotteries cannot be intentionally won. In this case, intuitively, Cindy did not intentionally or skillfully win the rigged lottery, nor did she intentionally pick the rigged ticket, even though believed truly and justifiably that buying the ticket handed to her would make her win the lottery. Here is another case of an uncontrolled action based on a Gettiered belief:

Barn Rendezvous Barney is involved in some hunting game. His clue tells him to stop at a barn on the side of a certain road and await further instructions. Unbeknownst to Barney, the road winds through fake barn county. Barney pauses at the first barn-like construction, but decides to move on, thinking that there's not enough shade for him to wait there comfortably. The next barn-like construction has just the right amount of shade, so Barney pulls over and waits. As luck would have it, he pulled over in front of the only real barn in fake barn county.

It is extremely natural to describe this case by saying that, while Barney succeeded to stop at a barn, his action was not under his control, for he did not even know that he was stopping at a barn

(since his belief that he was stopping at a barn is Gettiered). After all, he could have easily stopped at a fake barn, if only he had decided to wait in a sunny spot.⁸

Some might protest that we can concoct Gettier cases where intuitions pull the other way. Consider a variant of a case discussed by Setiya (2008), Setiya (2011) and recently discussed by Moss (2020). Jerry has been given a drug that paralyzes his hand. The doctor tells him that the drug will wear off at noon. Consequently, Jerry forms the intention to clench his fist at noon. It turns out that the doctor was looking at the wrong patient's charts. However, it also turns out that everything they said was true of Jerry. At noon, when Jerry successfully clenches his fist, it seems that his action can be intentional, despite being based on a Gettiered belief. In response, notice that in normal circumstances, someone can gain evidence that they are clenching their fist while they are doing so through proprioception. So, even if, before noon, Jerry did not know that he could clench his fist by intending to do so, it seems plausible that at noon, mid-clench, Jerry could know this. Of course, we can imagine a variant case where Jerry lacks this proprioceptive evidence, perhaps due to the drug's residual effects. But once we add this stipulation, it becomes less clear that Jerry's clenching is genuinely intentional.⁹

Some may worry that ETC can be challenged on other — arguably more direct — grounds. Consider Davidson's example of someone who is writing heavily on a page, intending to produce 10 legible carbon copies (Davidson (1978)). While writing, the agent does not know that they are producing 10 legible carbon copies. Still, if they succeed in their goal, it seems that they did so intentionally. As Piñeros Glasscock (2019) and others have noted, however, there is ground for resisting Davidson's counterexample, by observing that normally someone making ten carbon copies has the opportunity to check that they are succeeding (Thompson et al., 2011, p. 210). In such a case, the act of checking enables one to know that one is making ten carbon copies, and in checking, the agent would also manifest knowledge of how to perform the action. By contrast, if the carbon-copier does not bother to check what they are doing, it is much less clear that their

⁸For more examples of this sort, cf. Beddor & Pavese (2020) and Beddor & Pavese (2021).

⁹Beddor & Pavese (2020) discuss at some length some apparent counterexamples to intentional action requiring knowledge that have been proposed in the recent literature on know-how (especially Cath (2015), but also Poston (2009) and Carter & Pritchard (2015)) in which acting on a Gettiered belief allegedly can manifest one's know-how. We think that these examples are actually not genuine Gettier cases, as some recent experimental results also suggest (Carter et al. (2019)). A discussion of these cases would bring me too far afield.

action is intentional.¹⁰

Finally, consider skillful actions that have a significant chance of failure:

Home Run Babe Ruth comes up to bat and swings at the oncoming ball, intending to hit a home run. His bat connects with the ball and... Success! The ball soars over the outfield fence, exactly as planned.

Plausibly, Babe Ruth intentionally hit a home run. Indeed, we think the intuition that Babe Ruth's action is intentional is considerably stronger than the corresponding intuition in the carbon-copier case. But is ETC satisfied? Arguably, the answer is "No." After all, even the best baseball players in the world only hit a home run on a small fraction of the occasions in which they attempt to do so. So it is doubtful that Babe Ruth knew — when making contact with the ball — that he was hitting a home run. Taking a cue from Davidson's own remarks about the carbon-copier case and from Setiya's discussion of such cases, (Setiya, 2007, p. 25), one might replace ETC with Weakened ETC, which also defines controlled action in terms of knowledge:

Weakened ETC: Necessarily, an agent is in control of ϕ -ing, just in case they know that they are performing some action ψ in order to ϕ and they know this in virtue of being guided by their intention to ϕ and by their knowledge of how to ϕ .

While Babe Ruth does not know — mid-swing — that he is hitting a home run, he does know that he is performing some other action (namely, swinging the bat thus and so) in order to hit a home run. Indeed, **Weakened ETC** also helps with a different line objection, according to which habitual intentional action does not require knowledge that one is performing *that* action. Paul (2009) and Vekony et al. (2020) (among others) have argued that habitual and automatic action do not satisfy Anscombe's practical knowledge condition on intentional action. However, Weakened ETC remains unscathed by this challenge.

In conclusion, an epistemic theory of agentic control holds considerable promise: it can capture the insight of psychological theories of control — that at least *many cases* of intentional action require a monitoring condition — while at the same time overcoming the main counterexamples that afflict theories of agentic control in terms of psychological mechanisms, such as attention.

¹⁰See also Kneer (2021) for a yet different defense of Anscombe's account of intentional action against the challenge raised by Davidson's carbon copies example.

It is motivated by the commonsensical considerations that controlled action is a matter of knowing what one is doing while one is doing it; it can capture intuitions about controlled actions in a wide variety of cases; it is, further, motivated by theoretical conditions about the modal profile of agentive control. Finally, it is compatible with the primeness of agentive control, since it reduces agentive control to the holding of prime conditions.

8 From Control to Mentality

Agentive control is a feature of skillful actions that psychologists routinely attempt to explain. It is natural to expect that a general theory of controlled action would have to invoke the mental states of agents. Combining a prominent condition on intentional action and control — due to Anscombe (1963) — together with a knowledge-how condition, gives us a theory of agentive control that is sufficiently general to account for every case of controlled action. So a knowledge-based theory of agentive control is not incompatible with naturalistic approaches to agentive control in terms of attention and focus, of the sort provided by Wu (2016). Rather, it is a generalization of those approaches that is also extensionally adequate.

This argument from the epistemic theory of control to MENTALITY might seem to rely on the assumption that a theory of agentive control would have to appeal to the mental states of the subjects. Of course, this is not the only kind of theory of agentive control one could offer. For example, one might aim at a theory of agentive control that looks at how control is *implemented* — at the neural correlates of agentive control. A neurological theory of control of this sort would presumably not need to invoke knowledge. The point here is rather that *a* theory of agentive control can be given at the personal/representational level and that at this level it would have to involve the mental states of the subject. I have argued that agentive control is prime — so it cannot be decomposed into an internal and external condition. If it cannot be so decomposed, then the only adequate theory of agentive control is a theory that appeals to other prime conditions. The epistemic theory of control essentially invokes knowledge — both in the practical knowledge constraint and in the knowledge-how constraint. Thus, it reduces controlled action to the holding of some prime conditions. It can also capture the modal profile of controlled actions and a wide variety of judgments about intentional and skillful action.

This argument does not rely on drawing a line between distinctively psychological explanations of behavior and not distinctively psychological explanations of behavior. In fact, it differs from

Williamson's argument in that it does not depend on making a claim about folk psychological explanations of behavior at all. It also vindicates a more substantial role for knowledge in explaining actions than Williamson's argument does. According to the epistemic theory of agentic control, the role of knowledge in explaining actions is not confined solely to an explanation of resolute actions but extends pretty much to every skillful and intentional performance. My argument differs from Nagel (2011) in that it does not rely on moving from a claim about the concept of knowledge and its role in our explanatory practices to its mental reality. The epistemic theory of agentic control is not a theory about our practices of ascribing agentic control — it is a theory of a property that skillful behavior generally seems to possess and that psychologists try to explain. Finally, let me compare my argument to Gibbons (2001)'s argument for MENTALITY, which has come the closest to the argument developed in this paper. Gibbons (2001) argues that knowledge of the means are necessary for intentional action. However, Gibbons comes nowhere close to offering a theory of agentic control — he only argues for a knowledge condition on intentional action, nor does he embrace a practical knowledge condition on controlled action. Lacking a theory of agentic control, one might be skeptical about an argument going from a knowledge condition on intentional action to its mentality, for one might worry that there could be a more general condition that captures agentic control and of which knowledge is only a component. Indeed, it is easy to show that knowledge cannot provide all that is needed to overcome the problem of deviant causal chains, which psychological theories of agentic control were designed to solve. For example, consider the following variant of **Killing Nephew** — i.e., **Killing Knowledgeable Nephew**:

Killing Knowledgeable Nephew A nephew intends to murder his uncle. While driving one day, he sees a pedestrian that he recognizes as his uncle. So, the nephew plans to execute his intention and knows that if he accelerates the car, he will kill his uncle. But this reflection so unnerves him that he drives excessively fast and accidentally kills his uncle.

In this variant, the nephew knows that certain means — driving excessively fast — will kill the uncle. But this knowledge only deviantly causes him to kill his uncle, because his nervousness is also a factor. The action is still unintentional, whether or not it is caused by a belief or by knowledge. A plausible epistemic theory of agentic control ought to invoke a practical knowledge condition too, for only in this way it can offer a plausible generalization of the psychological theories considered in this paper and a full solution to the problem of deviant causal chains. Indeed,

it is worth noting that the epistemic theory of agentic control outlined here can capture deviant causal chains, including **Killing Knowledgeable Nephew**: for neither the killing nor the practical knowledge that he is killing his uncle happens in virtue of the nephew's intention or his knowledge of the means, so the action is neither controlled nor intentional.

9 Conclusions

While several prominent philosophers have argued for MENTALITY, many philosophers are still in disbelief. What are the theoretical benefits of thinking of knowledge as a mental state in its own right? This paper has argued that among the theoretical benefits of endorsing MENTALITY, there is a theory of agentic control that is appropriately general. I have argued that extant psychological theories of agentic control, in terms of monitoring mechanisms such as intention-attention coupling, fail to capture the full scope of agentic control. By contrast, an epistemic theory of agentic control can vindicate the important insights of these psychological theories, while also attaining the right scope. Appropriately formulated, an epistemic theory of agentic control can capture the modal profile of controlled action as well as overcome the problem of deviant causal chains by combining a prominent condition on intentional action (according to which, intentional action requires practical knowledge) with a knowledge-how condition. If this is correct, then it is when it comes to understanding the hallmarks of skillful performance that the theoretical benefits of thinking of knowledge as mental can be fully appreciated.

References

- Anscombe, G. E. M. (1963). *Intention*. Oxford: Blackwell.
- Baron-Cohen, S., Ring, H., Moriarty, J., Schmitz, B., Costa, D., & Ell, P. (1994). Recognition of mental state terms. *British Journal of Psychiatry*, *165*(5), 640–649.
- Bartsch, K., & Wellman, H. M. (1995). *Children talk about the mind*. Oxford University Press.
- Beddor, B., & Pavese, C. (2018). Modal virtue epistemology. *Philosophy and Phenomenological Research*, *45*(1), 1–19.
- Beddor, B., & Pavese, C. (2020). Skills as knowledge. *manuscript*.

- Beddor, B., & Pavese, C. (2021). Practical knowledge without luminosity. *manuscript*.
- Bratman, M. (1981). Intention and means-end reasoning. *The Philosophical Review*, *90*, 252–265.
- Call, J., & Tomasello, M. (2011). Does the chimpanzee have a theory of mind? 30 years later. *Human Nature and Self Design*, (pp. 83–96).
- Carter, A., & Pritchard, D. (2015). Knowledge-How and Epistemic Luck. *Noûs*, *49*(3), 440–453.
- Carter, J. A., Pritchard, D., & Shepherd, J. (2019). Knowledge-how, understanding-why and epistemic luck: an experimental study. *Review of Philosophy and Psychology*, *10*(4), 701–734.
- Cath, Y. (2015). Revisionary intellectualism and Gettier. *Philosophical Studies*, *172*(1), 7–27.
- Chisholm, R. (1966). *Freedom and Action*. New York: Random House.
- Christensen, W., Bicknell, K., McIlwain, D., & Sutton, J. (2015). The sense of agency and its role in strategic control for expert mountain bikers. *Psychology of Consciousness: Theory, Research, and Practice*, *2*(3), 340.
- Christensen, W., Sutton, J., & Bicknell, K. (2019). Memory systems and the control of skilled action. *Philosophical Psychology*, *32*(5), 692–718.
- Christensen, W., Sutton, J., & McIlwain, D. J. (2016). Cognition in skilled action: Meshed control and the varieties of skill experience. *Mind & Language*, *31*(1), 37–66.
- Davidson, D. (1978). Intending. In *Philosophy of History and Action*, (pp. 41–60). Springer.
- Davidson, D. (1980). Actions and events. *Clarendon Press, Oxford*, *3*, 1–41.
- De Villiers, J. (2007). The interface of language and theory of mind. *Lingua*, *117*(11), 1858–1878.
- Dreyfus, H. L. (2002). Intelligence without representation — Merleau-Ponty’s critique of mental representation and the relevance of phenomenology to scientific explanation. *Phenomenology and the cognitive sciences*, *1*(4), 367–383.
- Epley, N., & Waytz, A. G. (2009). Perspective taking. In *Encyclopedia of Human Relationships*, (pp. 1228–1231). Sage Publications, Inc.

- Fricker, E. (2009). Is knowing a state of mind? The case against. In P. in Greenough, & D. e. Pritchard (Eds.) *Williamson On Knowledge (Oxford: Oxford University Press)*.
- Fridland, E. (2014). They've lost control: Reflections on skill. *Synthese*, 191(12), 2729–2750.
- Fridland, E. (2021). The nature of skill: functions and control structures. In E. Fridland and C. Pavese, *The Routledge Handbook of Philosophy of Skill and Expertise*. Routledge, Oxon, (pp. 245–257).
- Frisch, M. (2020). Causation in Physics. In E. N. Zalta (Ed.) *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, fall 2020 ed.
- Gibbons, J. (2001). Knowledge in Action. *Philosophy and Phenomenological Research*, 62(3), 579–600.
- Ginet, C. (1990). *On Action*. Cambridge, Mass: Cambridge University Press.
- Goldman, A. (1970). *A Theory of Human Action*. Englewood Cliffs, New Jersey: Prentice Hall.
- Greco, D. (2016). Safety, Iteration, and Explanation. *Philosophical Issues*, 26(1), 187–208.
- Guttenplan, S. (1994). Belief, knowledge and the origins of content. *Dialectica*, 48(3-4), 287–305.
- Hampshire, S., & Hart, H. L. (1958). Decision, intention and certainty. *Mind*, 67(265), 1–12.
- Heyes, C. M. (1998). Theory of mind in nonhuman primates. *Behavioral and Brain Sciences*, 21(1), 101–114.
- Hogrefe, G.-J., Wimmer, H., & Perner, J. (1986). Ignorance versus false belief: A developmental lag in attribution of epistemic states. *Child Development*, (pp. 567–582).
- Hommel, B. (2000). The Prepared Reflex: Automaticity and Control in stimulus-response translation. *Control of Cognitive Processes*, (p. 247).
- Jackson, F. (2002). Critical Notice of 'Knowledge and its Limits' by Timothy Williamson. *Australasian Journal of Philosophy*, 80(4), 516–521.
- Keysar, B. (2007). Communication and miscommunication: The role of egocentric processes. *Intercultural Pragmatics*, 4(1), 71–84.

- Kneer, M. (2021). Success and knowledge in action: Saving Anscombe's account of intentionality. In T. Ciecierski, & P. Grabarczyk (Eds.) *Context Dependence in Language, Action, and Cognition*, (pp. 131–154). De Gruyter.
- Knobe, J. (2003). Intentional action and side effects in ordinary language. *Analysis*, 63(3), 190–194.
- Levy, Y. (2013). Intentional action first. *Australasian Journal of Philosophy*, 91(4), 705–718.
- Lewis, D. (1974). Causation. *The Journal of Philosophy*, 70(17), 556–567.
- Lin, S., Keysar, B., & Epley, N. (2010). Reflexively mindblind: Using theory of mind to interpret behavior requires effortful attention. *Journal of Experimental Social Psychology*, 46(3), 551–556.
- Loewer, B. (2012). Two accounts of laws and time. *Philosophical studies*, 160(1), 115–137.
- Magnus, P., & Cohen, J. (2003). Williamson on knowledge and psychological explanation. *Philosophical Studies*, 116(1), 37–52.
- McDowell, J. (1995). Knowledge and the internal. *Philosophy and Phenomenological Research*, 55(4), 877–893.
- McGlynn, A. (2014). *Knowledge first?*. Springer.
- McGlynn, A. (2017). Mindreading knowledge. *Knowledge First: Approaches in Epistemology and Mind*, (p. 72).
- Mele, A., & Moser, P. (1994). Intentional action. *Nous*, 28, 39–48.
- Mele, A., & Svedlik, S. (1996). Intention, intentional action and moral responsibility. *Philosophical Studies*, 82(3), 265–87.
- Miller, E. K. (2000). The prefrontal cortex and cognitive control. *Nature reviews neuroscience*, 1(1), 59–65.
- Molyneux, B. (2007). Primeness, internalism and explanatory generality. *Philosophical Studies*, 135(2), 255–277.

- Montero, B. (2010). Does bodily awareness interfere with highly skilled movement? *Inquiry*, 53(2), 105–122.
- Montero, B. G. (2016). *Thought in action: Expertise and the conscious mind*. Oxford University Press.
- Moss, S. (2018). *Probabilistic knowledge*. Oxford University Press.
- Moss, S. (2020). Replies to Edgington, Pavese, and Campbell-Moore and Konek. *Analysis*, 80(2), 356–370.
- Nagel, J. (2011). Knowledge as a mental state. *Oxford Studies In Epistemology*, 18(1), 16–25.
- Nagel, J. (2017). Factive and nonfactive mental state attribution. *Mind & Language*, 32(5), 525–544.
- Nagel, J. (forthcoming). The distinctive character of knowledge. *Behavioral and Brain Sciences*.
- Noonan, H. W. (1993). Object-dependent thoughts: A case of superficial necessity but deep contingency? In J. Heil, & M. A. (Eds.) *Mental Causation*. Oxford: Clarendon Press.
- Palmeri, T. J. (2006). Automaticity. *Encyclopedia of Cognitive Science*.
- Papafragou, A., Cassidy, K., & Gleitman, L. (2007). When we think about thinking: The acquisition of belief verbs. *Cognition*, 105(1), 125–165.
- Papineau, D. (2019). The disvalue of knowledge. *Synthese*, (pp. 1–22).
- Paul, S. K. (2009). *How we know what we're doing*. Ann Arbor, MI: Michigan Publishing, University of Michigan Library.
- Pavese, C. (2013). The unity and scope of knowledge. PhD Thesis Rutgers University.
- Pavese, C. (2015). Practical Senses. *Philosophers' Imprint*, 15(29), 1–25.
- Pavese, C. (2017). Know-how and gradability. *The Philosophical Review*, 126(3), 345–383.
- Pavese, C. (2018). Know-how, action, and luck. *Synthese*, (pp. 1–23).
- Pavese, C. (2020). Probabilistic knowledge in action. *Analysis*, 80(2), 342–356.

- Pavese, C. (2021). Knowledge, action, and defeasibility. In J. Brown, & M. Simion (Eds.) *Reasons, Justification, and Defeaters*. Oxford University Press.
- Peacocke, C. (1999). *Being known*. Clarendon Press.
- Piñeros Glasscock, J. S. (2019). Practical knowledge and luminosity. *Mind*, 129(516), 1237–1267.
- Posner, M. I., Snyder, C. R., & Solso, R. (2004). Attention and cognitive control. *Cognitive psychology: Key readings*, 205.
- Poston, T. (2009). Know how to be gettiered? *Philosophy and Phenomenological Research*, 79(3), 743–747.
- Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and brain sciences*, 1(4), 515–526.
- Prichard, H. A. (1950). *Knowing and believing*. Clarendon Press, Oxford.
- Ryle, G. (1949). *The Concept of Mind*. Chicago: Chicago University Press.
- Saxe, R. (2005). Against simulation: the argument from error. *Trends in cognitive sciences*, 9(4), 174–179.
- Schlosser, M. E. (2007). Basic deviance reconsidered. *Analysis*, 67(3), 186–194.
- Schlosser, M. E. (2008). Agent-causation and agential control. *Philosophical Explorations*, 11(1), 3–21.
- Setiya, K. (2007). *Reasons without rationalism*. Princeton University Press.
- Setiya, K. (2008). Practical knowledge. *Ethics*, 118(3), 3–44.
- Setiya, K. (2011). Knowing How. *Proceedings of the Aristotelian Society*, 112(3), 285–307.
- Shatz, M., Wellman, H. M., & Silber, S. (1983). The acquisition of mental verbs: A systematic investigation of the first reference to mental state. *Cognition*, 14(3), 301–321.
- Shepherd, J. (2014). The contours of control. *Philosophical Studies*, 170(3), 395–411.

- Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: Perceptual learning, automatic attending and a general theory. *Psychological Review*, 84(2), 127.
- Small, W. (2012). Practical knowledge and the structure of action. *Rethinking epistemology*, 2, 133–227.
- Smith, M. (2017). The cost of treating knowledge as a mental state. In A. Carter, E. Gordon, & B. Jarvis (Eds.) *Knowledge First, Approaches to Epistemology and Mind*, (pp. 95–112). Oxford University Press.
- Sodian, B., Thoermer, C., & Dietrich, N. (2006). Two-to four-year-old children's differentiation of knowing and guessing in a non-verbal task. *European Journal of Developmental Psychology*, 3(3), 222–237.
- Sosa, E. (1999). How to defeat opposition to Moore. *Nous*, 33.13, 141–153.
- Sosa, E. (2007). *A Virtue Epistemology*. Oxford: Clarendon Press.
- Sosa, E. (2009a). Knowing Full Well: the Normativity of Beliefs as Performances. *Philosophical Studies*, 142(1), 5–15.
- Sosa, E. (2009b). Timothy Williamson's 'Knowledge and its Limits'. *Williamson on knowledge*, (pp. 203–16).
- Stich, S. P. (1978). Autonomous psychology and the belief-desire thesis. *The Monist*, (pp. 573–591).
- Tardif, T., & Wellman, H. M. (2000). Acquisition of mental state language in Mandarin-and Cantonese-speaking children. *Developmental Psychology*, 36(1), 25.
- Thompson, M., Hornsby, J., & Stoutland, F. (2011). 7. Anscombe's intention and practical knowledge. In *Essays on Anscombe's intention*, (pp. 198–210). Harvard University Press.
- Toner, J., Montero, B. G., & Moran, A. (2015). Considering the role of cognitive control in expert performance. *Phenomenology and the Cognitive Sciences*, 14(4), 1127–1144.
- Vekony, R., Mele, A., & Rose, D. (2020). Intentional action without knowledge. *Synthese*, (pp. 1–13).

- Verbruggen, F., & McLaren, I. (2014). Banishing the control homunculi in studies of action control and behavior change. *Perspectives on Psychological Science*, 9(5), 497–524.
- Wellman, H. M., & Liu, D. (2004). Scaling of theory-of-mind tasks. *Child development*, 75(2), 523–541.
- Williamson, T. (1997). Knowledge as evidence. *Mind*, 106(424), 717–741.
- Williamson, T. (2000). *Knowledge and Its Limits*. Oxford, UK: Oxford University Press.
- Williamson, T. (2017). Acting on knowledge. *Knowledge first: Approaches in epistemology and mind*, (pp. 163–181).
- Wu, W. (2011). Confronting many-many problems: Attention and agentive control. *Noûs*, 45(1), 50–76.
- Wu, W. (2016). Experts and deviants: The story of agentive control. *Philosophy and Phenomenological Research*, 93(1), 101–126.