Ergo AN OPEN ACCESS JOURNAL OF PHILOSOPHY

Autonomy as Practical Understanding

REZA HADISI
University of Toronto

I offer a theory of autonomous agency that relies on the resources of a strongly cognitivist theory of intention and intentional action. On the proposed account, intentional action is a graded notion that is explained via the agent's degree of practical knowledge. In turn, autonomous agency is also a graded notion that is explained via the agent's degree of practical understanding. The resulting theory can synthesize insights from both the hierarchical and the cognitivist theories of autonomy with at least some aspects of the reason-responsiveness theories. Moreover, by treating practical knowledge and practical understanding as gradable notions, the paper offers a strategy to respond to enduring objections against cognitivism about intention, control, and autonomy.

Keywords: autonomy; strong cognitivism; practical understanding; intention; practical knowledge

In virtue of what is an action autonomous? I argue that an agent acts autonomously only when and because she acts from her *practical understanding* of what she is doing. In other words, autonomy is constituted by practical understanding.

First, a caveat. The thesis I aim to defend is a version of what we may call the *privileged activity* theories of autonomy. On this family of views, one's actions are autonomous only when and because they are attributable to her in virtue of a privileged kind of activity, e.g., because she has a second-order desire to perform the action, because she endorses it, because it follows from her policies and plans, etc.[1] My positive thesis identifies a new candidate for the privileged activity that makes actions autonomous: namely, practical understanding. But

1. Well-known examples of the view include Frankfurt (1971: 71; 2019), Watson (1975), Mele (1995), and Bratman (2005).

Contact: Reza Hadisi <r.hadisi@utoronto.ca>

https://doi.org/10.3998/ergo.6788

given the long and tired list of all the available options (higher-order desires, endorsement, valuations, plans, etc.), a reader may reasonably wonder: Why do we need yet another instance of the privileged activity theories of autonomy?

Autonomy is a kind of achievement for intentional actions. Consequently, our theories of autonomy cannot float free from our theories of intentional action. Almost anyone would accept this thought at this level of generality. However, arguably, the force of this point in the practice of theorizing about autonomy is not fully appreciated. In this paper, I will show how a familiar way of theorizing about autonomy takes a different shape once we fix our theory of intentional action in terms of what we can label as strong cognitivism about intention and intentional action. While strong cognitivism has been influential in action theory, its implications for theories of autonomy have not been much explored.[2] I will show that if intention is understood as identical with a kind of cognitive state, then intentional action, and consequently, autonomy, are to be understood as relevant kinds of cognitive achievements. This can result in a novel and interesting version of the privileged activity theories of autonomy.

The view that I advance is built around three commitments (§2.):

- **Volitional action** is constituted by the agent's *true practical belief* about what she is doing.
- **Intentional action** is constituted by the agent's *practical knowledge* of what she is doing.
- **Autonomous action** is constituted by the agent's *practical understanding* of what she is doing.

Thus, in this picture, we characterize different levels of agential achievement from merely volitional to fully autonomous action in terms of the levels of the agent's practical-cognitive achievements. I argue that by so doing, we can conceive of autonomy as an important agential achievement that is typically available to agents like us. As we will see, the resulting theory can nicely synthesize the insights of privileged activity theories of autonomy with at least some aspects of the reason-responsiveness theories of autonomy, especially as developed by Fischer and Ravizza (1998).

1. The Target Phenomenon

"Autonomy" is neither just a term of art (such as "supervenience"), nor merely a concept integral to ordinary moral thought and discourse (such as *responsibility*).

2. David Velleman (2007) is an important exception here. I will say more about the difference between my account and his view below.

As a result, it is advisable to start by setting out some desiderata for theories of autonomy by introducing constraints from our pre-theoretical normative intuitions about responsibility and self-government.

1.1 Two Desiderata for Theories of Autonomy

[1] Inclusivity

At the most general level, personal autonomy is a quality of an agent's activities,[3] namely the quality of instantiating *complete* self-government. To be sure, this definition is not very informative. Still, it helps us highlight a first desideratum for theories of autonomy by emphasizing that autonomy picks out only a proper subset of self-governing activities. In other words, any theory of autonomy must recognize that many self-governing activities are nonautonomous.

A tension concerning ordinary practices of moral blame and praise motivates this desideratum. One might be tempted to say that if an activity is not completely mine, then it would not be fitting to blame or praise me for it. The first desideratum is meant to guard us against this temptation. For instance, consider the following example:

> **Akratic[4] Terrorist:** Tony has been a member of a terrorist organization for a while. However, after careful deliberation, he has come to believe that, all things considered, he ought to leave the group and pursue civil engagement with his ideological opponents. He starts making arrangements, trying to find a new job, seeking new friends, etc. But when he receives an order to attack innocent civilians, he is swayed by his desire to hurt his opponents and acts against his all-things-considered judgment. He carries out a terrorist attack akratically.

3. I use *activity* in a broad sense, encompassing both productive and nonproductive kinds of doing. So, e.g., we can talk about both the activity of creating a statue (i.e., a productive activity) and the activity of remaining seated (i.e., nonproductive activity). The latter kind of activity is especially important for our purposes since it identifies *states* that can be constituted by *activities*. E.g., we can make sense of seating autonomously because although being seated is a state, sitting down and remaining seated are activities. Likewise, we can accept that knowledge and understanding are mental *states*, and yet hold that they can be constituted by autonomous activities, e.g., of coming to know and retaining one's knowledge. The distinction between productive and nonproductive activities can be traced back to Aristotle (340BCE/2019: 1.1.2, 1094a5). (I will, henceforth, use 'NE' to refer to Aristotle's *Nicomachean Ethics*.) The idea that we can apply this distinction to mental states like knowledge is defended by Boyle (2011).

4. Here, I am relying on a common-sense conception of "akrasia" as the broadly acknowledged experience of the weakness of will. Of course, the exact characterization of the phenomenon is philosophically contentious.

Tony's activity is, in some sense, self-governing—he carries out the terrorist attack not because he is forced to do it but through his own agency. However, the action is an instance of partial self-government because of its akratic character. Here, I am not assuming that complete self-government requires normative self-government, e.g., a type of self-government that is governed by norms of consistency. Rather, I am just noting that akratic self-government is that of a torn agent, and to that extent, her exercise of agency is partial. Akratic action is a paradigm example of partial self-government because, with respect to the activity in question, its agent is not fully in agreement with himself or herself.

Now, I think that commonsense delivers us the following verdicts. First, the *act* is attributable to Tony. Even if we accept that Tony's character has genuinely changed (in that he is not a committed terrorist anymore), all else being equal, it is still fitting to blame him for this particular action. More importantly, though, all else being equal, it would be fitting to criticize Tony's *character* for this action, too. To be sure, by stipulation, there is a positive change in his character in that he sincerely disavows his terrorist beliefs. After all, he is making arrangements to leave the organization. Nevertheless, he still shows deep character flaws. Indeed, we may say that the action is character-attributable to him, not despite his incontinence but in light of it. The hold of this inhumane anger on Tony and his incontinent relation to that anger shows us something deep about his character.[5]

If that is all on the right track, then according to what I call *Inclusivity desideratum*, our conception of autonomy should not render it as a necessary condition of mere act attributability of an activity nor a necessary condition for its character-attributability.[6] In particular, it is important to emphasize that a theory of autonomy is too restrictive if it cannot make sense of partial self-governing activities as activities that may ground appropriate deep blame or deep praise for action or character.[7]

[2] Typicality

But why should we accept that there is a difference between complete and partial self-governance to begin with? Aren't all my doings just my doing? This

5. The same point can be made if we reflect on cases of praising someone for doing the right thing akratically.

6. The basic idea behind this distinction between act and character attribution has to do with Aristotle's conception of virtue as a relatively stable 'state' (NE: Bk II). In the contemporary setting, the idea is reflected in the debate that started with Wolf (1993) and Watson (1996) concerning responsibility as attributability. Buss (1994; 2012) makes a similar observation about the relation between autonomy and attribution in a number of places.

7. I am assuming that the bar for radical revisionism about normative concepts must be very high. For a philosophical defense of this methodological assumption, see Nussbaum (2001: ch. 8).

might be especially worrying at this point since, so far, I have argued that the autonomy of an activity does not make a difference to its attributability in a zero-sum manner.

It is customary to motivate the idea that we have a special kind of relation to only a subset of our activities by thinking through contrasting examples. Frankfurt's willing and unwilling addict examples are perhaps the most familiar (1971: 12, 19).[8] However, I agree with Velleman (2015: 133) that those kinds of examples are too dramatic for capturing a phenomenon that is much more pervasive and familiar in the ordinary business of life. For better or worse, we can feel self-alienated: we find ourselves doing things or having attitudes (desires, beliefs, wonders, etc.) that we do not identify with. For example, you feel jealous of the accomplishments of a colleague although, otherwise, you value her as a friend. You ask her an uncharitable question at her talk, and then you see your action for what it is. You feel ashamed and weak. You think that your jealousy got the better of *you*. You wish you had more control over your own behaviors, feelings, and thoughts.

Now, I said that the experiences of self-alienation and the like are pervasive "for better or worse" because the mere fact that we can feel self-alienated does not entail that those feelings are always, or even ever, fitting. It may well be the case that we can feel self-alienated because we can be in bad faith.

We can accommodate such worries, at least to a degree, if we recognize a second desideratum for accounts of autonomy. To be sure, in order to make sense of autonomy as a special type of agency, we must assume that self-alienated agency is sometimes possible. However, we must avoid understanding the *specialness* of the relationship between an agent and her autonomous activities to mean that this kind of relation is *rarely* obtained. Indeed, as Kant forcefully argued (though the insight is surely not his alone), in the absence of overwhelming evidence to the contrary, there is a moral requirement to treat each other and ourselves under the assumption that our choices are autonomous.[9] But this default assumption would not be reasonable if we set the bar for autonomy too high. Given this normative commitment, we must construct our concepts of agency in such a way that self-deception, incontinence, and the like become the exception, not the rule.[10]

8. Example of this sort can already be found in Aristotle (NE: III.1. 1110a1-a25).

9. To be clear, it is a trademark of Kantianism to claim that respect for autonomy is a strictly necessary rational requirement, and that all moral obligations can be derived from this fundamental requirement of reason. But surely, other moral theories can account for a default and defeasible obligation to respect the autonomy in ourselves and others while rejecting the further Kantian ambitions. Cf. Korsgaard (2007)

10. There is certainly an empirical question as to whether certain types of agency are typical or atypical for us. However, the import of empirical facts for our theories of autonomy is not

We can thus formulate our second desideratum. Theories of autonomy must make sense of *typical* human activities as achieving some level of autonomy. I call this the *Typicality desideratum.*[11]

1.2 Summary

The reasons to adopt these two desiderata are primarily normative. I introduced the Inclusivity desideratum to make sense of the idea that instances of defective self-government are attributable to us in both senses of act attribution and character attribution. I introduce the Typicality desideratum because I am in broad agreement with Kantians that our default assumption must be that human activities are autonomous.

Now, in what follows, I offer a theory of autonomy that satisfies these desiderata. I do not claim that my theory is the only theory that can achieve this. However, my account satisfies these desiderata in a unique way. Despite their influence on action theory, contemporary cognitivist theories of action pioneered by Elizabeth Anscombe (1958) have said little about the notion of autonomy. Here, I rely on some resources that strong cognitivism about intention and intentional action offers us to account for the notion of autonomy. The resulting theory, I contend, will help us see the relation between autonomous and intentional action in a new light.

2. The Main Thesis

In this section, I argue that one can act autonomously only if and because one has practical understanding of what one is doing. But as I indicated at the outset, this account of autonomy relies on a strong cognitivist conception of intention and intentional action. So, I will start by clarifying what I mean by strong cognitivism. I do not aim to offer a defense of strong cognitivism in this limited space.

straightforward. For example, suppose someone assumes that any action that is not motivated by 'pure reason' is incontinent. Further, suppose it were an established empirical fact that human actions are typically motivated by emotions. Assuming that incontinence is a paradigm case of defective self-government, it would then follow that on this conception of incontinence, human actions are not typically autonomous. But, given that purported empirical fact, someone who holds this conception of incontinence cannot then also think that it is reasonable to have a default attitude that human actions are autonomous. In that case, what needs to be revised is presumably either one's conception of autonomy, or one's normative commitment to the value of autonomy.

11. To be clear, it is not obvious to me that Kant's own account of autonomy satisfies the Typicality desideratum. As an indication, Kant holds that we must remain skeptical as to whether *any* of our particular actions are genuinely autonomous (1785/1997: 4:407). If autonomy was a typical property of human activities, then it would be odd to express such a high level of skepticism.

Rather, I aim to show how the privileged activity theories of autonomy take a different shape if we start from this form of cognitivism about intention.

2.1 A Strongly Cognitivist Account of Intentional Action

Philosophers often talk about the intentionality of an action in an all-or-nothing manner. However, the folk psychological uses of the concept seem to admit of degrees. There are three kinds of evidence for this claim. First, psychological studies on attributions of intentionality purport to report a graded set of judgments (Phillips et al. 2015; Cusimano et al. 2019). Though, these studies do not always distinguish between degrees of confidence that an action is intentional vs. a judgment that an action is intentional to a degree. Second, publicly available linguistic data in English show that "intentional" can be used via scalar constructions and degree modifiers. Internet search results for "very intentional," "less intentional," or "quite intentional" suggest that these phrases are naturally used. Third, formal legal concepts also disclose graded uses of "intentional." For instance, Canadian law clarifies "willful" criminal action as denoting "a relatively high level of *mens rea*" (R. v. Docherty 1989). In the United States, first-degree murder is often described as "*completely* intentional" (ABC News 2016; Koeninger 2016; Bork 2021), and phrases such as "completely intentional" sexual assault (Cauterucci 2017) or "completely intentional" jury selection (Cineas 2021) are often used in media and news outlets.

Of course, there may be philosophical reasons to diverge from the folk psychological uses of terms like "autonomy," "responsibility," and "intentionality." However, I do not see any compelling reasons to be revisionary in this instance. As we will see, in my account, acknowledging the gradability of intention plays an important role in making sense of it in terms of degrees of cognitive success.[12]

Now, we can approach the gradability of intention from a relatively uncontroversial observation: an activity is intentional to the degree that it is under the agent's control. That is, I can have more or less control over what I do. By the same token, my actions can be more or less intentional. To echo Hursthouse's famous examples (1991: 58), when I hear that I have won the lottery and jump up and down in the spur of the moment, I will not exert as much control as when I deliberate and then resign from my job. Nevertheless, both seem to be instances of intentional action, the latter activity is just more intentional than the former. In short, actions are more or less intentional in virtue of the fact that we exert more or less control over them.

12. Other cognitivists have recently made similar observations about the gradability of intention (Marušić & Schwenkler 2018: § 3.1; Setiya 2008: 396; 2009: 129–131)

But how should we understand the relevant notion of control here? Strong cognitivism can be understood as a hypothesis that answers this question. More precisely, strong cognitivism about intention can be coupled with strong cognitivism about control to answer this question. On the strong cognitivist account, we must characterize an agent's control over her intentional action solely in terms of her practical cognition of what she is doing and how she is doing it.[13] The more an agent has practical cognition of what she is doing and how she is doing it, the more she is in control of what she is doing. By the same token, the more she has practical cognition of what she is doing and how she is doing it, the more intentional is her action.[14]

Note that control over action may not always be conscious or the result of an occurrent cognitive effort. For example, as many have noted (Paul 2009; Annas 2011: ch. 2), expert pianists or expert athletes have a high degree of control over their skilled behavior precisely because they do not need to be effortful in guiding their actions. Cognitivist theories of control can account for skilled and habitual control by identifying some habits as "intelligent habits," i.e., habits that reflect an agent's knowledge and understanding of what they are doing (Small 2021). In this picture, we can think of effortful control in intentional action as constituted by conscious practical cognition and knowledge-how, and intelligent habits and skillful control as constituted by implicit or dispositional instances of practical cognition and knowledge-how.[15]

As a working approximation, I characterize levels of practical *cognition* in terms of true practical belief, practical knowledge, and practical understanding. The first part of my proposal is then this. Assuming that the agent knows how to do something, we can say: an agent who merely has a true practical *belief* of what she is doing performs a minimally intentional action (i.e., merely voluntary). By contrast, an agent with practical *knowledge* of what she is doing performs a *fully* intentional action. I will unpack this latter claim first.

13. My thesis about control is a graded version of what Beddor and Pavese (2021) call the Epistemic Theory of Control. For alternatives to the cognitivist model of control, see Paul (2009), Shepherd (2021), and Wu (2023).

14. To be sure, this way of setting things up is in tension with a contemporary orthodoxy to distinguish between the *control* and *epistemic* conditions on intentional action (Fischer & Ravizza 1998: 11–12). As Fischer and Ravizza point out, the distinction can be traced back to Aristotle (NE: 1109b30-1111b5). I agree with Aristotle (and Fischer and Ravizza) that being coerced disqualifies an activity as a candidate for intentional action. For example, when the wind blows and moves someone's arms, then the person is not an agent at all because, as Aristotle puts it, she "contributes nothing." However, Fischer and Ravizza want to insist that even when there is no coercion, there is a distinction between the control that an agent has over what she does, and her epistemic relation to what she does. The cognitivist picture that I am urging in this paper is going to deny this latter distinction. On my account, when there is no coercion, what constitutes an agent's control over what she does is identical with her practical cognition of what she does. Cf. Mele (2010) who identifies the epistemic condition as a component of the control condition.

15. Beddor & Pavese (2021: 922, fn 2) make a similar suggestion.

One cautionary note: For the most part, I will focus on how practical cognition of what one is doing modulates the degree of the intentionality of an action. However, as we will see, we can get a similar correlation with knowledge-how, as well: fixing the agent's degree of knowing what they are doing, in performing an action, the more an agent knows how to do something, the more intentional their action is.[16]

In order to get a hold of the notion of practical cognition, I will start with a generic notion of a practical attitude and specify how the notion will be rendered in a strong cognitivist framework.

Take a *practical attitude* to be constituted by an agent's representation of what she is doing where the attitude in question is the cause of what the agent is doing. While there are many controversies about the exact nature of this causal relation, that our attitudes can be part of the causal explanation for our intentional actions should not be too controversial.[17]

One controversy that does arise after this point concerns the nature of the practical attitude of intention. According to cognitivists, intention (as a placeholder for *the* attitude that is the cause of the character of intentional action, that is, the cause which explains why an action is intentional under this-or-that description) is a practical attitude that can be assessed as veridical or non-veridical. Non-cognitivists deny this. That is, a cognitivist would claim that when, for example, I intentionally buy an ice cream for my son to make him happy, my intention is identical with an attitude that *(a)* causally explains my action, and *(b)* can be evaluated as veridical or non-veridical. Non-cognitivists accept (a), and thus identify intention with a *practical* attitude. However, they reject (b), and thus do not identify intention with a *cognitive* attitude. For example, desires are not evaluable in terms of veridicality. Hence, if desires wholly or partly constituted intentions, then strong cognitivism would be false.

As Sarah Paul (2009) points out, one could be a strong cognitivist in wholly identifying intention with a cognitive attitude, or one could be a weak cognitivist

16. Many cognitivists have argued that control requires practical knowledge of what one is doing, e.g. Anscombe (1958), Setiya (2008; 2009), Schwenkler (2015), and Velleman (2007). Others have emphasized that know-how must be part of the story of intentional action as well. For discussion, see Mele & Moser (1994), Small (2012), Pavese (2018), and Glasscock (2021).

17. Unfortunately, in this debate, different theorists use different conceptions of *causation.* The "causal theorists" often argue that the only relevant sort of causation is *efficient* causation (Davidson 1963). Some "non-causal theorists" argue that the only relevant type of causal relation between attitudes like belief about action and action is *formal* causation, e.g., Sarah Paul (2011) reads Anscombe along these lines. Some other "non-causal theorists" argue that the causal explanation in question is both efficient and formal, e.g., see John Schwenkler's (2019a: ch. 6) interpretation of Anscombe. But if we are open to this manner of speaking, then even teleological accounts of action are causal in that they speak of *final* causes. To avoid equivocating, I will specify the type of causation in question when necessary. On other occasions, when I want to remain neutral as to whether something is the formal, efficient, or final cause of an action, I will speak of the *cause of the character of an intentional action*. I will explain this notion below.

in maintaining that intention can be partly constituted by or even merely entail a cognitive attitude. As indicated earlier, I am trying to see how a privileged activity theory of autonomy takes shape under the assumptions of strong cognitivism. Thus, in the rest of this paper, I will use *cognitivism* as a substitute for "strong cognitivism."

As I said, my task here is not to defend cognitivism but to use it (or at least a version of it) to develop a theory of autonomous agency. In particular, my expository task in this section is this: I want to show how we can explain the graded notion of control that is constitutive of intentional action in cognitivist terms. Let's then focus on this specific task.

Recall that we want to explain a *fully* intentional action as constituted by an agent's practical knowledge of what they are doing and their knowledge-how. On the cognitivist account that I hold, my practical knowledge of what I am doing, along with the relevant know-hows, causally determine intentional action at least in the sense that they determine the *character* of intentional actions.

For instance, suppose that I am typing on the keyboard to write a paper in order to get famous. And let's say that my tapping on the keys is annoying my spouse. Now, consider the following cases[18]:

- **Unknowing-Annoyance**: I know that I am tapping on the keyboard as I am writing a paper, but I am ignorant of the fact that by so doing, I am annoying my spouse.
- **Knowing-Annoyance:** I know that I am tapping on the keyboard as I am writing a paper and that I am annoying my spouse by so doing.

It is relatively uncontroversial that in Unknowing-Annoyance, I am not intentionally annoying my spouse because I do not know that my typing annoys her. As Anscombe puts it, I can justifiably refuse an application of the "Why" question (Anscombe 1958: §7). But Knowing-Annoyance is ambiguous between two readings because there are at least two possible practical attitudes that are compatible with it:

- **Practical-Rep-1:** I know that I am tapping on the keyboard *as* I am writing a paper *in order to* get famous, and I know that I am tapping on the keyboard *in order to* annoy my spouse.
- **Practical-Rep-2:** I know that I am tapping on the keyboard *as* I am writing a paper *in order to* get famous. I know that tapping on the keyboard is annoying my spouse.

18. These cases resemble Bratman's (1991) purported counterexample to cognitivism about intention.

In Practical-Rep-2, I foresee that my typing causes annoyance in my spouse, but I do not represent the tapping as a means to that end (directly or indirectly). That is, given facts about what ends I have set for myself and what means I have adopted for those ends, it is not true that I am tapping on the keyboard *in order to* annoy my spouse. For instance, in this case, one could not use the principle of instrumental rationality to justifiably criticize me if my typing did not annoy my spouse and I did not seek other means to annoy her. They could not do so since, by supposition, I have not set the end of annoying my spouse for myself.

Note that at one level of description, Practical-Rep-1 and Practical-Rep-2 result in the same description of events. Since knowledge is factive, Knowing-Annoyance should be treated as describing the same happenings whether Practical-Rep-1 is my attitude or Practical-Rep-2: (1) I am typing, and (2) I am annoying my spouse by tapping on the keys. Indeed, Unknowing-Annoyance describes the same happenings, as well. However, because of my ignorance, Unknowing-Annoyance describes an intentional tapping on the keys and an unintentional act of annoying someone. There are, in turn, two readings of Knowing-Annoyance. When Knowing-Annoyance is read with Practical-Rep-1, I intentionally tap on the keys and annoy my spouse because of what I know about my adopted ends and means. When Knowing-Annoyance is read with Practical-Rep-2, I am intentionally tapping on the keys but merely foreseeing that I am annoying my spouse because of what I know about my adopted end and means. In other words, even when, in a sense, my attitudes do not change what happens in the world, my attitudes can determine the *character* of what happens in the world (i.e., which happenings are intentional under which description).[19]

In short, as Davidson (2001) has argued, actions are *intentional* only under some but not all descriptions that are true of them. In the cognitivist account that I am putting forward, we may now add that actions are under our control under some but not all true descriptions of them, as well. The subset of descriptions that are true of an action under which the action is intentional and under

19. To be sure, this does not address all the worries about explaining how a cognitive state (like belief or knowledge) can be causally efficacious. The worry is indeed deep enough that even someone like Velleman who is otherwise sympathetic to cognitivism finds some renderings of the theory to be "causally perverse" and "epistemically mysterious"(Velleman 2007: 103). So, there is, to be sure, a burden on cognitivists to offer clearer and less enigmatic explanations about the causal efficacy of beliefs and knowledge. However, since my aim is to show what a theory of autonomy would look like if we were to adopt strong cognitivism, I put aside this difficult question as something that any cognitivist would have to address eventually. After all, I think the explanatory burden of explaining the causal efficacy of cognitive states is shared between my view and any other form of strong cognitivism. For a promising cognitivist answer to this worry, see Schwenkler (2015; 2019b).

our control is at least partially demarcated by the practical attitudes of the agent of action.[20] Whatever else we may say about practical attitudes, I will assume that they are mental activities and states that can be employed to explain why an action is intentional under this but not another description.[21] (This is not yet to say that an action is intentional under a description *in virtue* of an agent's practical attitude. That will be argued later.)

Our account of intentional action is still incomplete because, as I indicated, it is crucial for my view that we should understand intentional action in terms of control, and control in terms of degrees of practical cognition. So far, I have only indicated how a fully intentional action may relate to a case of full practical knowledge. So, to get a grip on the notion of degrees of intentionality, we need to get a grip on the notion of degrees of practical cognition.

As Dretske (1988) teaches us, to grasp a cognitive notion of representation for an agent in a context, it would be crucial to understand what it means for her to *mis*represent in that context. Hence, I want to examine cases where the acting agent's practical attitude misrepresents. My proposal will have the following structure: the more I misrepresent what I am doing, the less control I exert over what I am doing. Hence, the more I misrepresent what I am doing, the less my action is intentional.

I will assume that, insofar as its content goes, an agent's intention when she is performing an intentional action is her representation of *what* she is doing: that is, a representation of an end (e.g., *why* are you boiling water?) and its means (e.g., *how* are you making lunch?), typically embedded in a network of means and ends (e.g., I'm making lunch in order to avoid writing, I'm avoiding writing in order to feel less stressed, etc.).[22] Additionally though, if intention is going to explain a particular action, then the representation that constitutes intention must encode, implicitly or explicitly, information about the particular action that the agent is performing. That includes information about who the agent is, what particular means she is adopting for what ends, the part/whole relations among these particularities (e.g., I am pressing this button to boil water as part of making tea to calm myself), the temporal order of sub-actions (e.g., I boil the water

20. As we will see, the contribution of knowledge-how will also be important in determining the intentional character of an action.

21. I am assuming that the explanation in question is causal *in some sense,* that is, in the sense of efficient, formal, and/or final causation. Of course, even this much is not completely uncontroversial. Cf. Goldman (1970) and Ginet (1990).

22. Two quick notes: first, the idea that the answer to "What are you doing?" is constituted by a series interrelated of "Why" and "How" questions that are embedded in other series of "Why" and "How" questions can be traced back to Anscombe (for discussion, see Small [2012: 162]). Second, as far as I can see, everything that follows could be said about intention about what the agent *will* do, as well. To keep things tidy, I will focus on the case of intention in the course of performing an action.

first, then I take the tea bags out, etc.), the locations of related obstacles, instruments, the agent's body, and so on.

Now, as Aristotle puts it, when someone acts intentionally, it is never the case that she would be ignorant of *all* these particularities.[23] To see this, consider an extreme case. Suppose that my "intention" radically misrepresents what I am doing. For example, I take myself to be making tea at home to calm myself, but thanks to a pill that I took with my lunch, I am hallucinating: In fact, I am throwing bananas at strangers in the park. In that case, in virtue of the fact that my "intending" attitude completely fails to determine the character of what I am doing, it also completely fails to represent what I am doing. I thus exert no control over what happens. By the same token, my activity is for the most part unintentional. Put differently, although I have an attitude that may be an efficient cause of what I am doing, the attitude is not the cause of the character of what I am doing under any description. Hence, to say that my action is not intentional is equivalent to saying that my attitude completely fails to represent what I am doing – which in turn is equivalent to saying that my attitude completely fails to cause the character of my action.

But, as usual, the less extreme cases are more interesting—i.e., cases where the agent exerts *some* degree of control and the action is somewhat intentional but not fully. In what follows, I consider three types of cases that have been used to argue against cognitivism: (1) cases where one's practical attitudes contain some error, (2) cases where an agent gets lucky in performing the action, and (3) cases of causal deviance. By going through these cases, I hope to explain the gradable notion of control and intentional action in cognitivist terms. Finally, I also hope to explain the grounding claim of cognitivism more clearly, i.e., the claim that an action is intentional *in virtue of* an agent's practical cognition.

First, consider the case of partial error. On the cognitivist picture that I am urging, an agent's practical cognitive error is inversely proportional to the agent's degree of control, and thus the degree to which what she does is intentional. We can illustrate this by thinking about Davidson's famous carbon copier, which was originally conceived as a counterexample to cognitivism:

> **[Partial-Error]** "[I]n writing heavily on this page I may be intending to produce ten legible carbon copies. I do not know, or believe with any

23. My thought in this paragraph is an adaption of what Aristotle says about involuntary action: "[T]he cause [of involuntary action] is ignorance of the particulars which the action consists in and is concerned with … . For an agent acts involuntarily if he is ignorant of one of these particulars. Presumably, then, it is not a bad idea to define these particulars, and say what they are, and how many. They are: who is doing it; what he is doing; about what or to what he is doing it; sometimes also what he is doing it with—with what instrument, for example; for what result, for example, safety; in what way, for example, gently or hard" (NE: 1111a1-5).

> confidence, that I am succeeding. But if I am producing ten legible carbon copies, I am certainly doing it intentionally" (Davidson 1980: 92).[24]

Here, we are urged to think that the agent intentionally makes ten carbon copies while she does not even believe she is doing so. Accordingly, it is claimed, her intention cannot be cashed out in terms of her knowledge.

Some cognitivists have tried to reply to this counterexample by rejecting that the carbon copier makes ten carbon copies intentionally because they reject the idea that she controls what she is doing (Thompson 2012: 210; Small 2012: 199). However, I think a better answer explains this case by relying on the gradability of control. As Davidson rightly puts it, in this case, the agent lacks at least one salient true belief about what she is doing, namely, she does not believe that she is making ten carbon copies. But as Davidson (2001: 50) notes elsewhere:

> Action does require that what the agent does is intentional under some description, and this, in turn requires […] that what the agent does is known to him under some description. But this condition is met by our examples.

I think cognitivists can and should welcome this part of Davidson's analysis.[25] The carbon copier surely has some true practical beliefs about what she is doing. Indeed, she has an abundance of true practical beliefs about what she is doing. For example, she believes that she is doing her job by *trying* to make ten carbon copies, she believes that she is using a pen to do it, she believes that she is pressing things on the table, she believes that this or that thing could be obstacles for her success, etc. Put differently, if the carbon copier were as deluded as our first example (the hallucinator in the park), then she would not be making ten carbon copies intentionally almost *at all*. However, her case is radically different. She does lack a salient belief about what she is doing. Still, she also has a large set of other relevant true practical beliefs, i.e., beliefs that successfully characterize her intentional action. Many of these practical beliefs are safe, i.e., it is not the case that the agent could easily be in circumstances where her belief would be false. While I will not assume a reductive relation between safe true beliefs and knowledge (Sosa 1999; Pritchard 2009; cf. Williamson 2000), we can at least say that a systematically interrelated set of safe true beliefs about something is a reliable indicator of knowledge about it. Hence, on the cognitivist proposal that I

24. For a similar case, see Bratman (2005: 38).

25. Beddor and Pavese express sympathy for a solution along the lines that I am going to propose. However, they do not consider the gradability aspect of this solution (Beddor & Pavese 2021: 922, fn 2).

want to assume, in virtue of all these other systematically related, safe, and true practical beliefs that the carbon copier has about what she is doing, we can say that she knows to a large extent what she is doing. Hence, her action is *to a large degree* intentional.

One may worry that it is one thing to *try* to make ten carbon copies intentionally, and another thing to *make* ten carbon copies intentionally. It is one thing to pick up a pen intentionally, and another thing to make ten carbon copies intentionally. And so on. Explaining the intentionality of the first set of actions (e.g., picking up the pen, placing the papers nearer to oneself, trying to make copies, etc.) does not explain the intentionality of the latter action (i.e., the successful performance of making ten carbon copies) at all.[26]

But I think that is much too strong. These are indeed different actions. Still, it does not follow that performing the first set of actions intentionally (e.g., the trying, the picking up the pen, etc.) does not explain the intentionality of the second action (i.e., making ten copies) "at all." The difference between trying to make ten carbon copies and making ten carbon copies is not the same as the difference between trying to make ten carbon copies and throwing bananas at strangers. That is so because, typically, trying to make ten carbon copies intentionally is part of making ten carbon copies intentionally. More generally, every course of action has, among other things, constitutive, typical, conceptual, instrumental, and temporal relations to other courses of action. On my account, an agent ϕs intentionally to the degree that she has practical knowledge of the acts that are constitutively, typically, conceptually, instrumentally, and temporally related to ϕ-ing.[27] So, e.g., if the carbon copier did not believe that she was even trying to make ten carbon copies, then her action would be even less intentional. Or, if the carbon copier did not believe that she was even picking up the pen, then her action is even less intentional. Insofar as these other practical beliefs are safe and true, we have reason to say that she is acting intentionally. However, her action is not *fully* intentional because she lacks a salient true practical belief: namely, she does not believe that she is succeeding at making ten carbon copies. Conversely, I think it is intuitive that Davidson's carbon copier would be making ten carbon copies even *more* intentionally if she knew she was making ten carbon copies. In short, one false practical belief about ϕ-ing does not suffice to render

26. Hence, other cognitivists have not denied that the carbon copier exerts control over these *other* actions. But, as I argue, the carbon copier's control over these other actions that are systematically related to making ten carbon copies implies a high degree (though, not full) control over making ten carbon copies.

27. Again, a reductivist could translate the degree of knowledge talk to degrees of safe true practical beliefs talk. But even if we are not inclined to analyze knowledge in that way, we could still think of systematically related safe and true practical beliefs as indicating practical knowledge.

my ϕ-ing unintentional because I may have many other true and safe practical beliefs about many other things that I am doing which are constitutively, typically, conceptually, instrumentally, and temporally related to ϕ-ing.[28]

Second, let's turn to counterexamples to cognitivism that exploit our intuitions about getting lucky. Consider an agent who succeeds in doing something by relying on a true practical belief that is not safe:

> **[Epistemic-Luck]** "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 it would be better to type in ten digits than 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 of his so doing that he "might thereby" shut down the reactor and prevent the explosion. What luck! He punched in the correct code, thereby preventing a nuclear explosion." (Mele & Moser 1994: 40)

It seems as if Fred does not know how to stop the disaster, and yet he manages to do so intentionally. For one thing, compare Fred with Mary, who, in the same situation, decides not to even try to enter a code, and instead makes a cup of tea and drinks it as the disaster unfolds. We want to say that Fred saves many lives intentionally, while Mary fails to do so. But, the critic argues, cognitivists seem hard-pressed to explain the intentionality of what Fred does. After all, he does not *know* the code.

In response, first, it is worth noting that Fred knows his end, he knows that his end is good, and he knows his end practically in that the end determines the intentional character of his action (i.e., in virtue of his knowledge of his end, his action is intentional under a description like "Fred is pressing the buttons *in order to* stop the disaster"). Indeed, his knowledge of his end, his knowledge of the goodness of his end, and the fact that this knowledge is *practical* (i.e., this is the attitude that characterizes his action) can be used to explain why we would judge

28. Of course, this is not just a numbers game. That is, just having 'more' true and safe practical beliefs does not amount to exerting more control over one's action under a description. Rather, what amounts to exerting more control is that I have more practical knowledge about activities that are constitutively, typically, conceptually, and instrumentally related to what I am doing. In other words, some elements of practical knowledge (or, some practical beliefs) are more *relevant* to my action under a given description. The more practical knowledge of the relevant sort I have, the more control I exert over what I do.

Fred differently from how we judge Mary even if he failed to stop the disaster. It is one thing not to know the means and another thing not to know the ends.[29]

Second, Fred certainly has some practical knowledge about the means, too. He knows *that* he should enter ten digits and not throw bananas at the computer. He knows *that* 1234567890 is a ten-digit code that he can use, but 1234 is not. He knows *how* to type. He knows that the chair beside the controller is an obstacle he needs to get out of the way first. And so on. So, when he gets lucky and enters the correct numbers, he acts intentionally to a degree, *not* in virtue of the lucky guess, but in virtue of his practical knowledge of his end, his practical knowledge of actions that constitutively, typically, conceptually, instrumentally, and temporally relate to successfully stopping the explosion, and his knowledge of how to perform these interrelated actions. If Fred knew the ten-digit code and stopped the reactor by employing that knowledge, then his action would be even more intentional. Or, if Fred did not know *how* to enter any digits but luckily fell with his fingers on the keyboard and pressed the right keys to stop the explosion, then his action would be even less intentional. In short, in the original example, Fred's action is *less* intentional than what it would be ideally because he is *less* in control of what he is doing – but he is less in control of what he is doing because he does not have *full* practical knowledge of the means to his end.[30]

Finally, let's see how my proposal would handle the cases of causal deviance. These cases are interesting because they put pressure on cognitivist theories from the opposite direction by showing that in some cases, the purported cognitive conditions may be in place, but the action may not be intentional (or under the agent's control).

> **[Deviant Causation]: "**A philosopher intends to knock over his glass in order to distract his commentator. However, his intention so upsets him that his hand shakes uncontrollably, striking the glass and knocking it

29. Once again, Aristotle has made this observation already. He argues that ignorance of the means of an action may be excusing, but ignorance of the ends is not: "For the cause of involuntary action is not … ignorance of the universal, since that is a cause for blame. Rather, the cause is ignorance of the particulars which the action consists in and is concerned with, since these allow both pity and pardon. For an agent acts involuntarily if he is ignorant of one of these particulars" (NE: 1110b30-1111a). I am suggesting that Aristotle's line can be used to at least partially explain the empirical results discussed in Nadelhoffer (2005; 2004) and Knobe (2003; 2004) about the imparity of judgments of blame and praise in similar cases.

30. Fred's case can show us another way in which our practical knowledge modulates our degree of control. We could imagine a case where Fred knows that he does not know the code. In that case, he may have second-order knowledge that he is acting not in a fully intentional manner because he knows what he is ignorant of is relevant to his complete control of what he is doing. This second-order knowledge could, in turn, give him some control over the activity of entering the ten digits. For example, he could intentionally wait until the last minute in case someone calls in and gives him the code.

> to the floor. Plainly, he does not intentionally knock over the glass, even though his intention results in the glass's crashing to the floor." (Adams & Mele 1989: 519)

In other words, in these cases, we seem to have a mere correlation between the agent's practical attitude (i.e., "I will knock over the glass in order to distract my commentator") and what happens in the world (i.e., he knocks over the glass and distracts his commentator). But this is a mere correlation in that what happens does not happen in virtue of the practical attitude. Hence, one might worry that practical knowledge does not constitute intention, even if it is often correlated with it.

In response, first, I want to resist the intuition that the philosopher in question knocks over the glass *completely* unintentionally. To see why, consider the following contrastive cases:

- **Elham-1** does not have an end to distract her commentator. However, she is so stressed by last night's election results that her hands shake uncontrollably, striking the glass and knocking it to the floor. As a result, she distracts her commentator.
- **Elham-2** does have an end to distract her commentator. However, she has a mistaken view about the means to that end, e.g., she thinks if she keeps quiet and listens to his comments, he will be distracted. However, the idea of sitting quietly stresses her so much that her hands shake uncontrollably, striking the glass and knocking it to the floor. As a result, she distracts her commentator.
- **Elham-3** knows full well that she wants to distract her commentator; she knows full well that knocking over the glass would do the trick (and she knows full well how to knock over a glass!). However, her intention so upsets her that her hand shakes uncontrollably, striking the glass and knocking it to the floor. As a result, she distracts her commentator.
- **Elham-4** knows full well that she wants to distract her commentator; she knows full well that knocking over the glass would do the trick (and she knows full well how to knock over a glass!). So, she knocks the glass to the floor to distract her commentator. As a result, she distracts her commentator.

As I see it, Elham's actions from 1 to 4 get increasingly more intentional. We could say Elham-1 knocks over the glass and distracts her commentator completely unintentionally because she does not even believe that she is doing anything in that vicinity. Elham-2 knocks over the glass and distracts her commentator somewhat intentionally because she had a true, though not very safe, belief that

she was doing so—in the end, she got very lucky and got it done. But I think it is a mistake to exploit our intuition about a case like Elham-1 or 2 and equate it with something like Elham-3. One simple way to see the difference is this: presumably, before her panic attack, Elham-3 pours the water into the glass with the intention of knocking it over and distracting the commentator; she orients her body towards the glass with that intention, she picks up the glass with that intention, etc. In other words, Elham-3 has practical knowledge of activities that are constitutively, typically, conceptually, instrumentally, and temporally related to knocking over a glass in order to distract someone, while Elham-1 and 2 lack most of the relevant practical knowledge.[31] In that sense, Elham-3's action of knocking over the glass to distract her commentator is more intentional because she has practical knowledge of all the other systematically related actions that she performs to that end.

But what exactly explains the difference between Elham-3 and Elham-4? To be sure, we could stipulate that there is an extended period when Elham-3 and Elham-4 have *the very same* set of practical attitudes: they both know full well what they are doing, and they both know how to do it, in such a way that it is characterizing the intentional character of what happens in the world. For instance, they both orient their bodies towards the glass in order to knock it over and distract their commentator, they both pour some water in the glass for that end, etc. However, when Elham-3's hand starts to shake uncontrollably, it is no longer the case that she has full practical knowledge of what she is doing or how she is doing it. That is so because, unlike desires, plans, and other non-cognitive states, the attribution of knowledge presupposes success. But presumably, Elham-3 does not represent the uncontrollable shaking of her hand as the means to knocking over the glass. She may know that she is uncontrollably shaking her hand *and* she is knocking over the glass. But this is just a theoretical knowledge of what is going on with her body. In other words, by stipulation of the purported counterexample, she does *not* have the following practical attitude: *I am shaking my hand in order to knock over the glass in order to distract my commentator*. Her practical attitude fails to amount to practical knowledge, and thus, she is not in control.[32]

I think it is important to note what exactly goes wrong with Elham-3's practical attitude, i.e., why exactly does her practical attitude fail to amount to practi-

31. Obviously, that is compatible with saying that Elham-1 has theoretical knowledge of all these possible ways of acting. In that case, Elham-1's knowledge is not practical in that it does not determine the character of what she does at all.

32. In explaining the difference between Elham-3 and Elham-4, cognitivists could also emphasize the role of knowledge-how (Beddor & Pavese 2021; Pavese 2021). One could say that Elham-4 knocks over the glass in virtue of her knowledge-how. However, Elham-3 does not act fully intentionally in knocking over the glass because her knowledge-how is not operant in causally explaining the shaking of her hand.

cal knowledge? After all, at least before her panic attack, she does hold a practical belief that *I will knock over the glass in order to distract my commentator.* And in a sense, one might say, that attitude turns out to be true.[33] However, I want to insist that this future-looking practical attitude does not amount to knowledge. Let me explain why.

In response, recall how we said that the content of an intention encodes information about who the agent is, what particular means she is adopting for what ends, the part/whole relations among these particularities, the temporal order of sub-actions, the location of related obstacles, instruments, her body, and so on. Of course, some of this information is encoded implicitly. Arguably, this means that some relevant information about causation is, at least implicitly, part of any intentional attitude as well. Now, when we form an attitude that by doing *x* we will cause *y* (e.g., by knocking over the glass, I will distract my commentator), we do so with an implicit *ceteris paribus* modifier. That is, we assume that knocking over the glass will cause the desired result "other things being right." Otherwise, our causal generalization will be false, and we recognize such causal claims as false.[34] Thus, Elham-3's practical attitude that *I am knocking over the glass in order to distract the commentator* turns out to be false because her belief about her own causal role in the world is based on a *ceteris paribus* modifier that is not satisfied. That is, since her hand is shaking uncontrollably, the conditions for the causal mechanism she was going to exploit to knock over the glass are not right. She thus lacks practical knowledge.

I hope this final point illustrates an advantage of the cognitivist theory of intention whereby the attribution of the intention (i.e., practical knowledge) does not depend on the agent's mental state alone. The agent *has* the intention insofar as her practical attitudes characterize what she does in the world correctly. In this sense, cognitivists can say that the cognitive aspect of intention is not merely an add-on when it comes to attributing intention. Rather, it is what constitutes the intention because, without the cognitive element, the attribution of intention to an agent could fly free from what the agent does in the world.

A final illustration of this last (and, as we will see later, very crucial) point about misrepresentation of one's own causal role. To borrow an example from Thomas Nagel, consider an agent who puts a coin in a pencil sharpener whenever she is thirsty and wants a can of soda (Nagel 1979: 39). There is evidently a correlation between her practical attitude and her kind of behavior. For example, if we knew somebody with this condition, and observed them putting a coin in

33. I say only "in a sense" because it does not become true that Elham-3 knocks over the glass *in order to* distract the commentator. What becomes true after the panic attack is that she knocks over the glass *and* she distracts the commentator.

34. What I say about causation echoes Cartwright's observation about causal laws and her rendering of *ceteris paribus* as "other things being right" (Cartwright 1980).

a pencil sharpener, we could reliably infer that she *wants* a can of soda. Does this agent act intentionally in putting a coin in the pencil sharpener in order to satisfy her thirst?

My answer, of course, has to do with grades of intentionality: *yes, she does act intentionally, but only to a degree*. This agent acts intentionally to the extent that she has practical knowledge of what she is doing and how to do it. She knows how to put a coin in a sharpener; she knows that she is putting a coin in a sharpener, etc. However, she does have a salient false practical belief about her causal role in the world: she thinks she is acquiring a can of soda by putting the coin in the sharpener. In virtue of this false practical belief, she is not fully in control of putting a coin in the sharpener *in order to* get a can of soda. And why do I insist that she is still exerting *some* control over her action and is acting somewhat intentionally? Compare this agent with someone who has the same condition but does not know how to put a coin in a pencil sharpener or mistakes bananas for coins, etc. Any such agent would be acting even less intentionally in virtue of their non-veridical practical attitudes.

To be in control, in other words, is a matter of mastering the world. This agent's lack of practical knowledge (i.e., her ignorance about what she *really* is doing) is an obstacle to her mastery of the world. In that sense, her lack of knowledge constitutes a partial lack of control, and by the same token, her action is less intentional.

So far, then, I have tried to draw a picture of intentional action where its constitutive control condition is reduced to a kind of epistemic condition. Hence, in this picture, an action is fully intentional when the agent has practical knowledge of what they are doing and they know how to do it. But knowledge of what one is doing is not reducible to knowledge of one simple proposition about one aspect of one's action. Our actions are embedded in networks of means and ends that are constitutively, typically, conceptually, instrumentally, and temporally related to one another. Hence, we exert more or less control, and our actions are more or less intentional depending on how much practical knowledge we have of the network of means and ends that characterize our actions. Somewhere on this spectrum, where the agent lacks many of the relevant true practical beliefs or their beliefs are by and large luckily true, we will have a case of merely voluntary action. On the other side of the spectrum, an agent acts fully intentionally when they have practical knowledge of *what* they are doing as embedded in the network of related means and ends, actions, intentions, etc. In short, then, intentional action is intentional to a degree that is determined by the degree of the agent's practical cognitive achievements.[35]

35. Many intentional actions take place in a multi-dimensional network of means and ends, and as a result, an action may be more intentional with respect to one set of ends and less intentional with respect to another. Put differently, at each point of time, we may have various, some-

2.2 Autonomy as Practical Understanding

Having construed the strongly cognitivist conception of intention, we can now see what a strongly cognitivist conception of autonomy should look like. Above, I characterized the difference between merely voluntary and intentional action in terms of a practical cognitive difference. Whereas a limited number of true practical beliefs about what I am doing renders my action voluntary, practical knowledge of what I am doing (as embedded in a network of means and ends) renders my action intentional. The next task is to show how practical understanding can be used to characterize autonomous action.

As I use the term, practical understanding is a kind of cognitive achievement where an agent veridically represents what they are doing as standing in a *dynamic* network of relations with their other actions, intentions, and values. The understanding in question is *practical* when it determines the *dispositional character* of what the agent does and how they do it. I propose that autonomy is constituted by practical understanding. Below, I unpack this claim by highlighting the difference between practical understanding and practical knowledge.

First, note that in formulating practical knowledge, we focused on the relation between a particular action and one's other (1) actions and (2) intentions. But in formulating practical understanding, we are interested in the relation between one's action and one's other (1) actions, (2) intentions, and (3) values.[36]

The underlying thought is this: one may act fully intentionally despite many of one's own values, but to act fully autonomously is to act in such a way that (among other things) reflects one's values as a whole.[37] To be sure, intentional-

---

times incompatible, plans and adopted ends. For example, I might be writing this paper because I have the end of discovering fundamental truths about human agency and because I want to get a better job. For the sake of argument, suppose these two ends are incompatible (e.g., because getting the better academic requires defending false views of a trendy philosopher). Strictly speaking, in these cases, there might not be a single 'spectrum' of degrees of intentionality. Rather, there may be distinct spectra: e.g., with respect to finding fundamental truths my action is less intentional because I am deluded about the causal efficacy of writing the paper, while with respect to finding a better job my action is fully intentional. However, for the purpose of this paper, I will simplify and speak of degrees of intentionality of an action as if the agent does not have competing or conflicting ends. Thanks to Mark Steen for pressing me on this point.

36. In a sense then, practical understanding presupposes practical knowledge. That is, because presumably, if I do not know my actions and intentions, then I would not be in a position to understand them in relation to my values. Hence, on my account, autonomy presupposes intentionality. Though, as I have laid things out, this is not an additive relation: just by making an action more intentional we do not end up with an autonomous action, because autonomy introduces a third element (values) and it asks for a different kind of relation (dynamicity).

37. This is *not* to assume the falsity of the guise of the good thesis about intentional action. One could go against most of their values while acting under the guise of the good as prescribed by even one opposing value.

ity requires a certain kind of coherence as well. Insofar as an agent sets an end for themself, and insofar as this end coheres with other actions that are constitutively, typically, conceptually, instrumentally, and temporally related to their end, intentional action requires some degree of local coherence among the agent's means and ends (e.g., I cannot fully intentionally boil water to make tea if I am also unplugging the tea kettle to plug in the toaster). That is so because knowledge requires some degree of local coherence: to use a toy example, I know that the Mississippi is a river only when I have some coherent true beliefs about what rivers are, where the Mississippi is, etc. But, in the present proposal, practical understanding goes beyond this local coherence into coherence with deeper facts about the agent and the outside world.[38]

In short, the first difference is this: through practical knowledge, our cognition determines the character of our action in relation to actions and intentions that are locally but systematically related to it. By contrast, through practical understanding, our cognition determines the character of our actions in relation to the actions, intentions, and *values* of the person as a whole. As we saw above, it is a central feature of cognitivism that the coherence of practical knowledge is not "purely structural," i.e., what is at stake is not just a coherence among one's mental attitudes. Practical knowledge requires coherence between one's practical attitudes and the facts about what one is doing in the world. As we will see, practical understanding will also require a similar kind of coherence that is not *merely* structural.

The second difference between practical understanding and practical knowledge has to do with the *dynamic nature* of practical understanding. Firstly, among epistemologists, there is almost a universal consensus that "connections" or "relations" are the *objects* of understanding (Grimm 2021). For example, I understand why Aisha is donating money to the UNHR only when I grasp the relation between this action and her other actions, intentions, and values. Or I understand Kant's categorical imperative only when I grasp the relation between the Formula of Humanity and the rest of what he says in the *Groundwork*. Secondly, and arguably, unlike knowledge, understanding requires more than a correct *static* representation of relations among different concrete or mental objects. Understanding also requires a grasp of the dependence relations that make these connections *dynamic* (Gopnik et al. 2004).

38. Practical understanding is a "higher" cognitive achievement in that it requires cognizing more things (i.e., how my action relates to my values as well as intentions) and cognizing the relation among its objects dynamically. However, I am not suggesting that it is always all-things-considered *better* to act from practical understanding than to act from practical knowledge. Arguably, we could be in situations where all-things-considered it is better to fully intentionally do something that does not fit well with one's deeper values and commitments. On my proposal, such actions will be less autonomous than they could be, but it does not follow that all-things-considered they will be less good.

For example, suppose that I am making tea intentionally. That is, I have practical knowledge of what I am doing in that I have practical knowledge of boiling some water, taking an herbal teabag out, taking a break from writing, etc., and I know how to do these. Arguably, I can possess this knowledge as long as I know the *actual* relations between my intentions and actions. Moreover, this knowledge is practical in that it determines what, in fact, is true in the world: i.e., my knowledge determines what intentional descriptions of my action are true. To be sure, I may falsely believe that I am boiling water while, in fact, I am boiling vodka. In that case, my practical belief does not determine *whether* vodka is being boiled. However, my practical belief does determine whether vodka is being boiled *intentionally* or not.

Contrast my knowledge of how my actions and intentions are actually related to one another to my understanding of the dynamics of dependence relations among my actions, intentions, and values. That is, suppose that in addition to that knowledge, I can further grasp how these relations *could* and *should* be given different reconfigurations. For example, suppose that I am making tea intentionally. That is, I have practical knowledge of what I am doing: I know a network of interrelated actions and intentions. But now, to grasp the dynamic dependence relations among these actions requires having a grasp of how things could and should go if I changed my ends, means, intentions, and values. In other words, not only does practical understanding require a deeper coherence (i.e., coherence not only with related actions and intentions but also coherence with one's values), but it also requires dynamicity or responsiveness to reasons. An agent has practical understanding of what they are doing insofar as they act in such a way that is responsive to dependence relations between their values, intentions, and other actions in a dynamic manner: i.e., they respond to changes, they respond to conflicting and reinforcing relations, etc. Behaviorally, this kind of disposition to respond to reason can be reflected in patterns of persistence, replanning, managing resources, etc. For example, suppose that I value my mental health, and I understand that having tea more often instead of watching the news would improve my mental health. This understanding becomes practical when I act in such a way that *if* I ran out of tea, I *would* allocate more time and energy to go to the store and buy tea instead of watching the news on the couch. When my actions are characterized by these dispositions, i.e., when my understanding of the dynamics between my actions, intentions, and values are reflected in my dispositions to respond to reasons in my environment, I act autonomously.

In short, understanding what I am doing requires grasping my action as embedded in the network of my values and other intentions and actions in a way that veridically reflects relations of dependence in that network. Moreover, this understanding is *practical* in that it determines dispositional facts about how

I respond to dynamic changes in my environment in light of my values and projects: i.e., my understanding determines whether and how I respond to reasons that come from the relation between what I am doing and my values and other actions and intentions. And once again, this is clearly a graded notion: I could be doing something more or less autonomously in that my action could be responsive to the dynamics of my values, and other intentions and actions with more or less sensitivity and vigor. I could make tea in a way that is dynamically responsive to my commitment to not overwork myself, but not in such a way that is responsive to my valuing of my mental health. In that case, I'm acting less autonomously than I could.[39]

Hopefully, what I said above sheds some light on what I mean by practical understanding as distinct from practical knowledge. But why should we think that autonomy is constituted by practical understanding?

I noted at the outset that my account is a version of the privileged activity theories of autonomy. We can think of this family of theories as taking the following shape:

> **Privileged Activity Theories:** an agent *S* performs an action ϕ autonomously in virtue of a privileged mental activity *A* that establishes a special relation between ϕ-ing and *S*.

But as many have noted, this kind of theory might be subject to a kind of regress. For instance, as Buss puts it, "If I must do something in order to govern my own agency, then must I not also do something in order to govern this self-governing activity?" (2012, 656). In our case, if I must have practical understanding in order to govern my own agency, then must I not also have practical understanding in order to govern my practical understanding? I think an efficient way to see why

39. Velleman defines intentional and autonomous action by reference to self-knowledge and self-understanding as well (2007: ch. 2, ch. 6; 1985). Velleman's important work certainly inspires my proposal, but I revise it in at least two significant ways. First, unlike Velleman, I do not rely on the idea that intentional and autonomous actions constitutively aim at self-knowledge and self-understanding. To be sure, it is compatible with my account that one's practical knowledge and practical understanding would have implications for one's self-knowledge and self-understanding. But that is evidently different from Velleman's strong (and, controversial) commitments about the aim of intentional and autonomous action. Second, and relatedly, for Velleman, intentional action requires *aiming* at self-knowledge, and autonomous action requires *aiming* at self-understanding. However, on my account, intentional action requires (successful) practical knowledge, and autonomy requires (successful) practical understanding. If I understand Velleman correctly, intentionality and autonomy are features of agency that reflect the 'quality' of our will. On my account, however, intentionality and autonomy are features of agency that reflect the degree of our cognitive *achievements*, which in turn, reflect the degree to which our activities are self-governing. (Many thanks to one of the reviewers and the area editor for helping me see the difference between my own view and Velleman's view more clearly).

practical understanding is a good candidate for the privileged activity that constitutes autonomy is to see how my proposal can handle this objection.

In response, I want to consider what it means to be an agent who has practical understanding of what they are doing but is still non-autonomous. Presumably, we are asked to consider someone who has acquired her practical understanding in a non-autonomous way, e.g., someone who takes a pill and suddenly acquires a grasp of dynamic relations between what she is doing and her other actions, intentions, and values. Or perhaps better: we are asked to consider someone who acquires her practical understanding as a result of brainwashing or intensive propaganda. Are we forced to accept these cases as possible, and if so, do these show that practical understanding is not constitutive of autonomy?

I will focus on the case of brainwashing, as I worry that taking a pill might amount to a case of mere physical coercion.[40] Moreover, the brainwashing example will also help me to show that my version of the privileged activity theory of autonomy is especially friendly towards the reasons-responsive theories, which are often portrayed as theoretical rivals.

So, can my practical understanding of what I am doing be the result of brainwashing? If so, should we call the actions of a brainwashed person autonomous?

First, consider a "good" case of brainwashing. Mary is "brainwashed" by a group of scientists and philosophers who feed Mary only accurate information about the world. Mary has access to no other sources of information, and so, all her beliefs are formed on the basis of facts that these scientists provide her. Based on this information (which includes information about means/end relations, about what she is doing, and, let's say, even about what is valuable to her), Mary gains both practical knowledge and practical understanding of what she is doing. That is, the knowledge and the understanding are given to her through this positive brainwashing regiment. She then acts in such a way that her knowledge and her understanding characterize her actions and dispositions. Can we say that she is acting intentionally and autonomously?

40. Are these pills coercive in the same sense that a gust of wind may move my arm and hit the person next to me? If so, then these cases are not very interesting. Note that the official thesis of the paper is that practical understanding is necessary for and constitutive of autonomy. It makes no claim to sufficiency. In general, constitutive conditions are necessary but not sufficient because they leave room for possible defeaters. For example, plausibly, following the rules of chess is constitutive of playing chess, but it is not sufficient for it. For instance, I could show you how a game of chess unfolded by moving the white pieces and telling you how to move the black pieces. We won't be playing a game of chess, although we would be following the constitutive rules of the game. In a sense then, we can restate the thesis by saying: an agent acts autonomously only if and because of her practical understanding of what she is doing, provided that certain undermining conditions do not hold (e.g., external coercion). From a very different perspective, also see Carter (2020) for an argument for why the 'pill-taking' examples do not apply to knowledge and similar states.

There is certainly an element of environmental luck involved here, but that is not a reason to conclude that she does not act intentionally or autonomously. To be sure, typically for creatures like us, knowledge and understanding do require the capacity to sift through reliable and unreliable information. However, the example does not show that Mary lacks the capacity to sift through reliable and unreliable information. Mary does not need to often rely on this capacity because her sources of information just happen to be very reliable. Mary thus knows what she is doing, and she understands it. On my account, then, whenever this knowledge and understanding characterize her actions and dispositions to respond to reasons, she acts intentionally and autonomously.

Put differently, the "good" case of brainwashing does not amount to a loss of agency because it is not a genuine case of brainwashing after all. Of course, we could detail the thought-experiment in a variety of ways, but a plausible way of redescribing the case is this: our scientists facilitate an excellent 'teaching' environment for Mary and reliably provide her with facts about the environment and herself.[41] Arguably, that is the case because attribution of learning, and thus successful knowing and understanding does not presuppose the possibility of failing to know and failing to understand. Hence, the fact that the scientists filter out false claims for Mary does nothing to her ability to learn in this environment. In other words, Mary's situation does not undermine her knowledge and understanding because there is no principle of alternate possibilities that applies to these cognitive achievements. And if my account of intention and autonomy is on the right track, that is the case because knowledge and understanding constitute agency; they are not things that we do with our agency.[42]

Now, compare the above case with the more familiar brainwashing cases where someone is fed bad information. Suppose Tommy has been 'brainwashed' with racist propaganda since childhood. Among other things, this should mean the following: Tommy has some false beliefs about the relation between a racialized group and certain features of the world. Now, suppose that he performs the following action: he donates a big chunk of money to a racist organization to help them fight the racialized group. Is Tommy acting intentionally? Is he acting autonomously?

The main lessons that I want to draw from this second example are the following: (1) It is very likely, though not necessary, that Tommy's action is not fully intentional. (2) By the same token, it is very likely, though not necessary,

41. Thanks to Mark Steen for helping me clarify this point.

42. Pamela Hieronymi (2008) makes a similar point in the defending the notion of responsibility for belief in the absence of doxastic voluntary control. The underlying thought behind this paragraph is that the possibility of failure is not a necessary condition for the possibility of agency. Pettit and Smith make a similar point about freedom to believe and to desire (1996: 444).

that Tommy's action is not fully autonomous. However, as we will see, although my view is compatible with what I will call moralism about autonomy, I do not make these claims by assuming that autonomous action is necessarily moral. On my account, Tommy's action is likely to be only partially intentional and only partially autonomous because it is likely that he misrepresents the world and misrepresents what he is doing. I believe that unpacking these two lessons will help us see the relation between the reasons-responsiveness theories of autonomy and my account, as well.

First, recall the earlier observation that part of the content of a practical attitude is a belief about one's own causal efficacy. I noted that such misrepresentations are inversely proportional to one's intentionality because one could lose touch with reality and thus lose mastery of the world by acting on the basis of a false causal belief. Now, it is very likely that the brainwashed Tommy is not acting *fully* intentionally nor fully autonomously because he has some false practical beliefs about his causal role in the world. For example, he believes that by donating money to the racist group, he is helping eradicate an immigrant group that "hurt the economy." To be sure, there might be no conceptual necessity that racist actions always depend on such factual mistakes. By the same token, then, my account does not predict that evil actions are necessarily not fully intentional or fully autonomous (at least not for the reason of getting the facts about one's own causal role wrong).

Second, and importantly, this is not to let Tommy off the hook too easily, either. Donating money to a racist organization to fight the racialized group is not done in the void: there are intentions and actions that are constitutively, typically, conceptually, instrumentally, and temporally related to it. Tommy has practical knowledge of his action, and thus his action is intentional to the extent that he knows these related intentions and actions, and insofar as this knowledge determines the character of his action. For example, he is intentionally paying some money to a group that promotes violence and is not making any factual mistakes in this case. We can thus say that he donates to the racist organization to fight the racialized group somewhat *intentionally,* not in virtue of his false practical belief that by so doing he is helping the economy, but because, fully intentionally, he is promoting violence (assuming that the promotion of violence against racialized minorities is systematically related to racist causes). Moreover, he promotes violence fully intentionally because he has practical knowledge of many actions constitutively, typically, conceptually, instrumentally, and temporally related to promoting violence.

Finally, and this will take us to the crux of my proposal: as stated earlier, the theory of autonomy as practical understanding is compatible with what I call moralism about autonomy, i.e., the view that an action is fully autonomous only if it is moral. It is so because someone might be tempted by the idea that

moral facts are like any other ordinary fact around us, and that we know them in much the same way as we know other facts about the world. Certain forms of reductive naturalists may even be tempted to say that misrepresentation of a moral fact is on par with misrepresentation of a causal claim, e.g., a causal claim about the promotion of wellbeing. With that kind of view, any case of acting on the basis of moral ignorance would be only partially intentional and partially autonomous for the same reasons as what we just saw above.

However, that sort of view is not our only option. In my proposal, autonomy requires practical understanding. Given how I have construed the notion of practical understanding as understanding the relation between one's actions, intentions, and values, it does follow that if one is mistaken about what one is doing or what one is intending, one is not acting fully autonomously. In other words, already in this picture, we can see that autonomy is not achieved by a mere meshing of what I *think* I am doing and my values. Rather, it requires grasping a relation between what I *know* I am doing and my values. But there is an ambiguity left: Is autonomy, in this picture, achieved by grasping the relation between what I know I am doing and (1) what I *know* to be valuable *to me*, or (2) what I *know* to be valuable *simpliciter*?[43]

Obviously, if we go with option (2), we are back to some sort of moralism about autonomy. But option (1) is still available and attractive as well. Importantly, even taking option (1) does not render the theory of autonomy as practical understanding into a merely structural view of autonomy. Let me explain this last point.

According to a theory that is often portrayed as a rival to the privileged activity theories of autonomy, to say that an agent is autonomous is to say that she is reasons-responsive. Fischer and Ravizza, for instance, argue that reasons-receptivity is constitutive of autonomy and characterize it by saying that it "involves a pattern of actual and hypothetical recognition of reasons (some of which are moral reasons) that is understandable by some appropriate external observer. And the pattern must be at least minimally grounded in reality" (1998: 90). However, Fischer and Ravizza present their theory in contrast to the privileged activity theories (which they call 'mesh theories') (1998: 186). They write, "We believe that the problem with *all* mesh theories, no matter how they are refined, is that they are purely structural and ahistorical" (1998: 186).

To say that a theory of autonomy is ahistorical is to say that in determining whether an action is autonomous, it only looks at the time-slice state of the agent at the time of the action. For example, a second-order desire theory of autonomy is ahistorical in that it does not ask any questions about the history and etiology

43. As noted earlier, this is on the assumption that I can understand the relation between *p* and *q* only if I know *p* and I know *q*. Hence, I can understand the relationship between my actions, intentions, and values only if I know them.

of an agent's second-order desires at the time of the action. This is problematic because, the intuition goes, someone who has been brainwashed to have a second-order desire in favor of something may not be exercising full autonomy.

Likewise, to say that a theory of action is purely structural is to say that in determining whether an action is autonomous, it only looks at coherence relations between an agent's attitudes. For example, seemingly, I could wholeheartedly identify with all my desires and beliefs, while I have many false beliefs. And if autonomy is constituted by wholehearted identification, it looks as if I could be so deeply mistaken about the object of my desire and yet be autonomous just because my beliefs and desires mesh well together.

Now, my account of autonomy as practical understanding is neither purely ahistorical nor purely structural. To that extent, my account is a version of the privileged activity theories of autonomy that accommodates some insights of the reasons-responsiveness theories.

It is easier to see why the theory is not purely structural. Practical understanding, as I have defined it, concerns the relationship between one's intentions, actions, and values. To be sure, since I want to remain neutral as to whether what is at stake is knowing what is valuable simpliciter vs. knowing what is valuable to me, the theory could be interpreted as more or less structural. But in neither case will the theory be purely structural. On this account, my understanding of what I am doing does not fly free from (i) what I am actually doing in the world, (ii) what my actual intentions are, nor from (iii) how my actions and intentions would, as a matter of fact, interact with what is valuable.[44]

For example, I may mistakenly hold that, given how much I value *x*, if I cannot do *a* to promote *x*, then I must allocate energy and resources to do *b* because I represent *b* as promoting *x* as well. This representation purports to be a case of practical understanding, and as such, it may determine my dispositions in action. However, this representation would fail to be a case of practical understanding if I am mistaken that *b* promotes the value *x* (be it *my* value or what is actually valuable). Hence, given what my practical understanding is *about* (i.e., about how my actions and intentions, as a matter of fact, may interact with my values), autonomy as practical understanding asks for much more than a mere structural mesh among my attitudes. My representation of the dynamic relations between actions, intentions, and values amounts to practical understanding, and thus autonomy, only when I get the dynamic relations right. (Again, note that this kind of factual misrepresentation could happen whether we construe the notion in terms of *my* values or what is valuable *simpliciter*).

We said that practical knowledge is constitutive of intention because it reflects an agent's mastery of the world as they try to affect it. Likewise, under-

44. Valuable *simpliciter* or to me, depending on which option we take.

standing the relevant dynamic relations of what one is doing is a way of mastering the world. It is a way of mastering the world in relation to our actions and intentions so that we can act on our values. But we would not be mastering the world to act on our values if we were getting things wrong about which actions would promote our values. Thus, in the same way, the cognitive character of autonomy is not a mere afterthought. Autonomy is constituted by cognition because it is in virtue of cognition of these dynamic relations that we can exercise complete self-government.

In short, practical understanding requires that my mental states fit together and *at least* fit with what I am doing in the world. This is one way in which cognitivism about intention pays off when it underpins our theory of autonomy and makes a privileged activity theory friendlier to the insights of reasons-responsiveness theories. Autonomy as practical understanding is not purely structural.

But also, the theory of autonomy as practical understanding is not purely ahistorical. This point is less clear and perhaps harder to defend. At an abstract level, practical understanding is not purely ahistorical because practical understanding concerns not just a static coherence between my current actions, intentions, and values. Rather, practical understanding requires a dynamic coherence: that is, it requires that my actions and intentions respond fittingly with how my values could and should change. And arguably, this kind of dynamicity is a matter of getting things right *over time*: that is, it is a matter of getting things right about my own mental states over time and a matter of getting things right about what I do in the world over time. More concretely, as we saw with the cases of brainwashing, the theory can remain sensitive to *how* an agent gets to a mental state because it is not concerned with the mere match between a higher-order desire and a lower-order desire, a plan and sub-plans, etc. But it is also concerned with getting things right and forming the right dispositions in relation to them.

In short, depending on how we interpret the theory, autonomy as practical understanding offers us a version of privileged activity theories of autonomy that is at least a close relative of reasons-responsiveness theories of autonomy.

Let me close by briefly surveying the desiderata that I mentioned at the outset. Recall our two desiderata: Inclusivity and Typicality. On the proposed account of autonomy as practical understanding, one may act intentionally (with practical knowledge of what they are doing) and yet be nonautonomous (without practical understanding of what they are doing). In that sense, inclusivity is satisfied. Put differently, inclusivity is satisfied because practical knowledge does not entail practical understanding.

Moreover, inclusivity is also satisfied in another sense, namely that at least one version of the theory is not committed to moralism about autonomy. That is, if we think of practical understanding as concerned with an agent's knowledge of what is valuable to her, then the agent's immoral action may be fully

autonomous. But even on the interpretation that what is at stake for practical understanding is knowing what is actually valuable, my proposal can still make sense of immoral actions as autonomous *to a large degree.* This was illustrated in the example of Tommy and his racist acts.

Finally, according to strong cognitivism, when we act with a high degree of intentionality, we have a high degree of practical knowledge of our action as embedded in a network of other actions and commitments. Although there may be no necessary conceptual relation between *knowing* this network and *understanding* the dynamics of those relations with our values, I think it is reasonable to say that, typically, our knowledge of such complexities is accompanied by a degree of understanding. This explains why our intentional actions are normally autonomous, and thus, Typicality is also satisfied.

3. Conclusion

I have argued that a strong form of cognitivism about intention can be a basis for a novel and interesting version of the privileged activity view of autonomy. This type of privileged activity theory can acknowledge and accept certain insights from the reasons-responsiveness theories of autonomy as well. At every step of the way, the theory relied on a graded notion of practical knowledge and understanding, coupled with a graded notion of intentionality and autonomy.

But putting aside my cognitivist project, I believe that my strategy of taking seriously the idea that intentionality and autonomy are graded notions could also be repurposed for some other privileged activity theories of autonomy. I suspect that doing so would help the privileged activity theories to overcome the suspicion that these theories cannot make sense of nonautonomous activities as activities of the agent at all (i.e., they fail Inclusivity), or that they turn autonomous activity into a rare or idealized sort of activity (i.e., they fail Typicality).[45] My final (unargued) proposal is that if intentionality and autonomy are systematically treated as graded notions, many other privileged activity theories could also make sense of our nonautonomous actions as truly ours and our agency as typically autonomous.

Acknowledgments

Thanks to Sara Aronowitz, Sarah Buss, Corey Cusimano, Michael McKenna, Alex Silk, and Mark Steen for helping me with this paper.* Many thanks to the

45. For example, see Buss (2012: 656–57).

* Correction notice: at the time of publication, Sarah Buss was mistakenly omitted from the list of acknowledgments. This has been corrected as of 11/5/2024.

audience at the *Epistemic Autonomy in Kant and Beyond* workshop at the Instituto de Filosofía UC (especially since my presentation was not about Kant!). Finally, thanks to the area editor and the two anonymous referees at *Ergo* for very helpful and constructive comments and suggestions.

References

ABC NEWS (November 1, 2016). Jurors in Ex-Cop Ray Tensing's Murder Trial Visit Scene of Fatal Cincinnati Shooting. Accessed August 10, 2023. https://abcnews.go.com/US/jurors-cop-ray-tensings-murder-visit-scene-fatal/story?id=43213659

Adams, Frederick and Alfred Mele (1989). The Role of Intention in Intentional Action. *Canadian Journal of Philosophy*, *19*(4), 511–531. https://doi.org/10.1080/00455091.1989.10716780

Annas, Julia (2011). *Intelligent Virtue* (1st ed.). Oxford University Press.

Anscombe, G. E. M. (1958). *Intention* (2nd ed). Harvard University Press.

Aristotle (c. 340 BCE/2019). *Nicomachean Ethics*. Terence Irwin (Trans.). Third Edition. Hackett.

Beddor, Bob and Carlotta Pavese (2021). Practical Knowledge without Luminosity. *Mind*, *131*(523), 917–934.

Bork, Cory (2021, March 31). George Floyd's Killer, Derek Chauvin, Should Face Harsher Charges. *The Fordham Ram* (blog). Accessed August 10, 2023. https://the-fordhamram.com/80187/opinion/george-floyds-killer-derek-chauvin-should-face-harsher-charges/

Boyle, Matthew (2011). 'Making Up Your Mind' and the Activity of Reason. *Philosophers' Imprint*, *11*(17).

Bratman, Michael (2005). Planning Agency, Autonomous Agency. In James Stacey Taylor (Ed.), *Personal Autonomy: New Essays on Personal Autonomy and Its Role in Contemporary Moral Philosophy* (33–57). Cambridge University Press.

Bratman, Michael (1991). Cognitivism About Practical Reason (Review of Practical Reflection, by J. David Velleman). *Ethics*, *102*(1). 117–128.

Buss, Sarah (1994). Autonomy Reconsidered. *Midwest Studies in Philosophy*, *19*(1), 95–121. https://doi.org/10.1111/j.1475-4975.1994.tb00281.x

Buss, Sarah (2012). Autonomous Action: Self-Determination in the Passive Mode. *Ethics*, *122*(4), 647–691. https://doi.org/10.1086/666328

Carter, J. Adam (2020). Epistemic Autonomy and Externalism. In Kirk Lougheed and Jonathan Matheson (Eds.), *Epistemic Autonomy* (21–40). Routledge.

Cartwright, Nancy (1980). The Truth Doesn't Explain Much. *American Philosophical Quarterly*, *17*(2), 159–163.

Cauterucci, Christina (2017). Jurors In Taylor Swift's Sexual Assault Case Had to Admit Whether They Were Fans of Her Music. *Slate*. Accessed, August 8, 2017. https://slate.com/human-interest/2017/08/jurors-in-taylor-swifts-sexual-assault-case-had-to-admit-whether-or-not-they-were-fans.html

Cineas, Fabiola (2021). Lawyers Left Racism out of the Trial over Ahmaud Arbery's Death. Here's Why. *Vox*. Accessed November 24, 2021. https://www.vox.com/22801394/racism-ahmaud-arbery-murder-trial

Cusimano, Corey, and Geoffrey P. Goodwin (2019). Lay Beliefs about the Controllability of Everyday Mental States. *Journal of Experimental Psychology: General*, *148*(10), 1701–1732. https://doi.org/10.1037/xge0000547

Davidson, Donald (1963). Actions, Reasons, and Causes. *Journal of Philosophy*, *60*(23), 685. https://doi.org/10.2307/2023177

Davidson, Donald (1980). Intending. In *Essays on Action and Events* (83–102). Clarendon Press.

Davidson, Donald (2001). Agency. In *Essays on Actions and Events* (43–62). Oxford University Press. https://doi.org/10.1093/0199246270.003.0003

Dretske, Fred (1988). *Explaining Behavior: Reasons in a World of Causes*. The MIT Press.

Fischer, John Martin and Mark Ravizza (1998). *Responsibility and Control: A Theory of Moral Responsibility*. Cambridge University Press.

Frankfurt, Harry (1971). Freedom of the Will and the Concept of a Person. *Journal of Philosophy*, *68*(1), 5–20. https://doi.org/10.2307/2024717

Frankfurt, Harry (2019). Identification and Wholeheartedness. In John Martin Fischer and Mark Ravizza (Eds.), *Perspectives on Moral Responsibility* (170–187). Cornell University Press. https://doi.org/10.7591/9781501721564-007

Ginet, Carl (1990). *On Action*. Cambridge University Press.

Glasscock, Juan S. Piñeros (2021). The puzzle of learning by doing and the gradability of knowledge-how. *Philosophy and Phenomenological Research*, *105*(3), 619–637.

Goldman, Alvin (1970). *A Theory of Human Action*. Princeton University Press.

Gopnik, Alison, Clark Glymour, David M. Sobel, Laura E. Schulz, Tamar Kushnir, and David Danks (2004). A Theory of Causal Learning in Children: Causal Maps and Bayes Nets. *Psychological Review*, *111*(1), 3–32. https://doi.org/10.1037/0033-295x.111.1.3

Grimm, Stephen (2021). Understanding. In Edward N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy*. Accessed August 10, 2023. https://plato.stanford.edu/archives/sum2021/entries/understanding/

Hieronymi, Pamela (2008). Responsibility for believing. *Synthese*, *161*(3), 357–373.

Hursthouse, Rosalind (1991). Arational Actions. *The Journal of Philosophy*, *88*(2), 57–68. https://doi.org/10.2307/2026906

Kant, Immanuel (1785/1997). *Groundwork of the Metaphysics of Morals*. Ed. and Trans. by Mary J. Gregor. Cambridge University Press.

Knobe, Joshua (2003). Intentional Action in Folk Psychology: An Experimental Investigation. *Philosophical Psychology*, *16*(2), 309–325. https://doi.org/10.1080/09515080307771

Knobe, Joshua (2004). Intention, Intentional Action and Moral Considerations. *Analysis* *64*(2), 181–187. https://doi.org/10.1093/analys/64.2.181

Koeninger, Kevin (2016, November 1). Campus Cop on Trial for Killing Cincinnati Driver. Accessed August 10, 2023. https://www.courthousenews.com/campus-cop-on-trial-for-killing-cincinnati-driver/

Korsgaard, Christine (2007). What's Wrong with Lying? In Jonathan Eric Adler and Catherine Z. Elgin (Eds.), *Philosophical Inquiry: Classic and Contemporary Readings* (577–587). Hackett Publishing.

Marušić, Berislav and John Schwenkler (2018). Intending Is Believing: A Defense of Strong Cognitivism. *Analytic Philosophy*, *59*(3), 309–340. https://doi.org/10.1111/phib.12133

Mele, Alfred (2010). Moral Responsibility for Actions: Epistemic and Freedom Conditions. *Philosophical Explorations*, *13*(2), 101–111. https://doi.org/10.1080/13869790903494556

Mele, Alfred R. (1995). *Autonomous Agents: From Self-Control to Autonomy*. Oxford University Press.
Mele, Alfred R. and Paul K. Moser (1994). Intentional Action. *Noûs*, *28*(1), 39–68. https://doi.org/10.2307/2215919
Nadelhoffer, Thomas (2004). The Butler Problem Revisited. *Analysis*, *64*(3), 277–284. https://doi.org/10.1093/analys/64.3.277
Nadelhoffer, Thomas (2005). Skill, Luck, Control, and Intentional Action. *Philosophical Psychology*, *18*(3), 341–352. https://doi.org/10.1080/09515080500177309
Nagel, Thomas (1979). *The Possibility of Altruism*. Princeton University Press.
Nussbaum, Martha (2001). *The Fragility of Goodness: Luck and Ethics in Greek Tragedy and Philosophy*. Cambridge University Press. https://doi.org/10.1017/CBO9780511817915
Paul, Sarah K. (2009). How We Know What We're Doing. *Philosophers' Imprint*, *9*(11), 1–24.
Paul, Sarah K. (2011). Deviant Formal Causation. *Journal of Ethics and Social Philosophy*, *5*(3), 1–24. https://doi.org/10.26556/jesp.v5i3.55
Pavese, Carlotta (2018). Know-how, action, and luck. *Synthese*, *198*(Suppl 7), 1595–1617. https://doi.org/10.1007/s11229-018-1823-7
Pavese, Carlotta (2021). Knowledge and mentality. *Philosophical Perspectives*, *35*(1), 359–382. https://doi.org/10.1111/phpe.12150
Pettit, Philip, and Michael Smith (1996). Freedom in Belief and Desire. J*ournal of Philosophy*, *93*(9), 429–449. https://doi.org/10.2307/2940892
Phillips, Jonathan, Jamie B. Luguri, and Joshua Knobe (2015). Unifying Morality's Influence on Non-Moral Judgments: The Relevance of Alternative Possibilities. *Cognition*, *145*(December), 30–42. https://doi.org/10.1016/j.cognition.2015.08.001
Pritchard, Duncan (2009). Safety-Based Epistemology: Whither Now? *Journal of Philosophical Research*, *34*(July), 33–45. https://doi.org/10.5840/jpr_2009_2
R. v. Docherty (1989). 2 SCR. 941. *Supreme Court of Canada*. Accessed August 10, 2023. https://decisions.scc-csc.ca/scc-csc/scc-csc/en/item/528/index.do
Schwenkler, John (2015). Understanding 'Practical Knowledge.' *Philosophers' Imprint* *15*(15). http://hdl.handle.net/2027/spo.3521354.0015.015
Schwenkler, John (2019a). *Anscombe's Intention: A Guide*. Oxford Guides to Philosophy. Oxford University Press.
Schwenkler, John (2019b). Practical Knowledge. In John Schwenkler, *Anscombe's Intention: A Guide* (155–200). Oxford University Press. https://doi.org/10.1093/oso/9780190052027.003.0006
Setiya, Kieran (2008). Practical Knowledge. *Ethics*, *118*(3), 388–409. https://doi.org/10.1086/528781
Setiya, Kieran (2009). Practical Knowledge Revisited. *Ethics*, *120*(1), 128–137. https://doi.org/10.1086/606000
Shepherd, Joshua (2021). *The Shape of Agency: Control, Action, Skill, Knowledge*. Oxford University Press.
Small, Will (2012). Practical Knowledge and the Structure of Action. In Günter Abel and James Conant (Eds.), *Rethinking Epistemology* (Vol. 2, 133–227). De Gruyter.
Small, Will (2021). The Intelligence of Virtue and Skill. *The Journal of Value Inquiry*, *55*(2), 229–249. https://doi.org/10.1007/s10790-021-09813-1
Sosa, Ernest (1999). How Must Knowledge Be Modally Related to What Is Known? *Philosophical Topics*, *26*(1/2), 373–384.
Thompson, Michael (2012). *Life and Action: Elementary Structures of Practice and Practical Thought*. Harvard University Press.

Velleman, J. David (1985). Practical Reflection. *The Philosophical Review*, *94*(1), 33–61. https://doi.org/10.2307/2184714
Velleman, J. David (2007). *Practical Reflection*. Center for the Study of Language and Information.
Velleman, J. David (2015). What Happens When Someone Acts? In *The Possibility of Practical Reason* (123–143). Michigan Publishing Services.
Watson, Gary (1975). Free Agency. *Journal of Philosophy*, *72*(8), 205–220. https://doi.org/10.2307/2024703
Watson, Gary (1996). Two Faces of Responsibility. *Philosophical Topics*, *24*(2), 227–248. https://doi.org/10.5840/philtopics199624222
Williamson, Timothy (2000). *Knowledge and Its Limits*. Oxford University Press.
Wolf, Susan (1993). *Freedom Within Reason*. Oxford University Press.
Wu, Wayne (2023). *Movements of the Mind: A Theory of Attention, Intention and Action*. Oxford University Press.