Acting on knowledge-how

Timothy Williamson[1]

Abstract

The paper explains how to integrate the knowledge-first approach to epistemology with the intellectualist thesis that knowing-how is a kind of knowing-that, with emphasis on their role in practical reasoning. One component of this integration is a belief-based account of desire.

Knowledge and its Limits (KAIL, Williamson 2000) and 'Knowing How' (KH, Stanley and Williamson 2001) were written at roughly the same time. Although KAIL's knowledge-first approach is quite consistent with KH's intellectualism, I did not explicitly integrate them into a unified account—even though KAIL highlights connections between knowledge-that and action, while the point of KH is to subsume knowledge-how under knowledge-that. This paper explains in more detail how intellectualism about knowledge-how fits into the knowledge-first programme, especially with respect to practical reasoning.

KAIL connects knowledge and action in several ways. First, it emphasizes the autonomous role of knowledge in the explanation of action. One acts on what one knows, and sometimes states of knowing explain action better than do the corresponding states of believing. Second, it treats believing as a state whose functional norm is knowing, and so believing without knowing as defective; since acting on what one believes but does not know is acting on defective beliefs, such actions are defective in origin. Thus there is a knowledge norm for practical reasoning (see also Hawthorne and Stanley 2008). Third, KAIL connects knowledge and action in a different way, by drawing an analogy between them. In doing so, it reacts to the Davidsonian paradigm of belief-desire psychology, where rational actions are explained by combinations of beliefs and desires (Davidson 1980). Since KAIL takes knowledge rather than belief as its starting-point, it naturally raises the question: what stands to

---

✉ Timothy Williamson
  Timothy.williamson@philosophy.ox.ac.uk

[1] University of Oxford (New College), OX1 3BN Oxford, U.K.

desire as knowledge stands to belief? If one understands the contrast between belief and desire in Anscombe's terms as the difference between mind-to-world and world-to-mind *directions of fit*, one might ask to similar effect: if knowledge is the good case for the mind-to-world direction of fit, what is the good case for the world-to-mind direction of fit?[1] KAIL's answer to both questions is: action (where 'action' means *intentional action*, as it will throughout this paper). The analogy helps motivate the knowledge-first approach: marginalizing knowledge in theoretical philosophy is as wrong-headed as marginalizing action in practical philosophy.

More recently, I have reworked the analogy between knowledge and action (Williamson 2017, 2018). I realized that starting with the question 'What stands to desire as knowledge stands to belief?' concedes too much to the Davidsonian picture, by letting it set the terms of the analogy. If action stands to desire as knowledge stands to belief, then action is, in a natural sense, *closer* to desire than to belief. But that undermines the traditional (and Davidsonian) idea that belief and desire are the twin inputs to practical deliberation, playing roughly symmetrical roles in determining action. Instead, we should start in the right place, and ask: what stands to action as belief stands to knowledge? Both knowledge and belief are on the input side of practical deliberation: knowledge in the good case, the case where the input came as it should, belief in both good and bad cases, whether the input came as it should or not. The answer to our new question should be, like action and unlike desire, on the output side of practical deliberation, and, like belief and desire, present in both good and bad cases, whether the output went as it should or not. More specifically, just as a belief can in principle figure as an immediate input to practical deliberation, irrespective of what led to it, we seek something that can in principle figure as an immediate output from practical deliberation, irrespective of what it led to. The natural candidate is *intention*. Practical deliberation outputs an intention to φ; in the good case, one intentionally φ's, while in the bad case one does not. More specifically, an intention can in principle figure as an immediate output from practical deliberation, irrespective of what it led to. In that respect, intending to φ is a better candidate than *trying* to φ, because the latter often requires a more extended sequence of actions: someone can intend to run a marathon before he dies without ever trying to run a marathon before he dies. In short, intention stands to action as belief stands to knowledge.

The revised analogy between knowledge and action displaces desire. But desire must still be in the picture *somewhere*. In the schematic terms above, desire has the same profile as belief: it is on the input side of practical reasoning, and it is not confined to a good case. Thus a natural hypothesis is that in some sense desires *are* beliefs, or at least *consist in* beliefs. Obviously, the thought is not that whenever one desires that P, one believes *that P*; rather, it is that whenever one desires that P, one believes something else concerning the proposition that P, perhaps that *it would be good that P*. But, even if desires are beliefs, it does not follow that 'desire' can be rigorously defined in terms of 'belief', any more than, since desires are mental states, it follows that 'desire' can be rigorously defined in terms of 'mental state'.

---

[1] See Anscombe 1957: 57–58 and, for discussion, Humberstone 1992 and Frost 2014 (and references therein).

Of course, treating desires as based on belief is deeply controversial (see Stocker 1979, Lewis 1988, 1996, Velleman 1992; contrast Tenenbaum 2007). But Sect. 1 explains one way of developing it into a defensible and explanatory view. Section 2 uses that view to build a schematic account of practical deliberation that integrates a knowledge-first approach with intellectualism about knowledge-how. Section 3 briefly hints at ways of deepening the picture.

1 Desires and beliefs

Some evidence that desires consist in beliefs is that desires can be *mistaken*. A vegetarian may regard herself as having previously had a mistaken desire to eat meat. A lapsed Christian may regard himself as having previously had a mistaken desire to go to church. If we ask them what mistake they had been making, the vegetarian might say: 'I used to think that eating meat was a good thing to do, but it wasn't'; the lapsed Christian might say: 'I used to think that going to church was a good thing to do, but it wasn't'. Thus the vegetarian treats her previous desire to eat meat like a false belief about the goodness of eating meat; the lapsed Christian treats his previous desire to go to church like a false belief about the goodness of going to church. 'Good' here need not be understood in a specifically moral sense. Presumably, the vegetarian never thought that eating meat was *morally* good; more likely, she used to think that it was enjoyable, nourishing, morally neutral, and good all-things-considered. Many desires concern non-moral goods. One can selfishly look after one's own good, and 'Evil be thou my good' is no contradiction.

Mistaken desires are not confined to reflective agents. Iris (the flowering plant) is poisonous to sheep, but they still want to eat it when they see it. Their desire to eat it is properly described as *mistaken*. The sheep's false belief can be put into words as something like 'That's good to eat', obviously in a non-moral sense of 'good'.

A more theoretical reason for supposing that desires consist in beliefs is that doing so helps us articulate practical deliberation from the agent's perspective as a potentially rational thought process. An agent desires that P and believes that if she does A, she will bring it about that P. No conflict between doing A and any of her other desires or beliefs obtrudes on her thinking. She rationally decides to do A. What does the agent herself think in this process? Her means-end belief figures as the premise 'If I do A, I'll bring it about that P'. But how does her desire that P figure? It cannot figure as the premise 'P': that would be the merest wishful thinking, and in any case would pre-empt the need for practical reasoning about how to bring it about that P. Can it figure as the premise 'I desire that P'? That proposal risks over-intellectualizing. Non-linguistic animals and young children *have* desires, which move them to action without being represented by the agent *as* desires. The proposal also misrepresents the motivation of normal agents, even self-aware ones. Faced with a drowning child in a pond, someone who needs the premise 'I want her not to drown' in his practical reasoning before he decides to save her is pathologically self-regarding; it's not about him, it's about her. A better premise would be 'She mustn't drown'.

An alternative proposal is that 'P' by itself figures in the thought process, though somehow *as desired* rather than *as believed*. What that means is unclear. The cogency

of ordinary reasoning depends on truth-preserving connections of content between the premises and the conclusion.[2] The conclusion *follows from* the premises, or at least is *probable conditional on* them. Such connections do not depend on whether the premises are believed, or only supposed—or desired. Our chances of understanding practical reasoning are far better if we can subsume the relevant standard of validity under a more general kind of truth-preservation from premises to conclusion than if we must concoct some new sort of validity that somehow depends on whether the premises are marked 'believed' or 'desired'.

A salient example of an approach that faces difficulties in reconstructing first-personal practical reasoning is *decision theory*. Its calculations of maximal expected utility in terms of the subjective probabilities and utilities of a supposedly rational agent do not correspond to any natural reasoning that the agent herself might go through. For instance, if her credence (subjective probability) in a proposition p is 0.5, what is the corresponding premise for her to start reasoning from? She will assume neither p nor its negation. She might assume 'My credence in p is 0.5', but that is a proposition about her own psychological state, and fails to capture the world-directedness of normal practical deliberation. It is a doxastic analogue of a premise of the form 'My subjective utility for the child's not drowning is x' in the pathologically self-regarding agent's practical reasoning about whether to save the child. To base one's decision entirely on premises about one's own psychological states is a bizarrely solipsistic approach to practical and moral life. It also raises the problem that one may not know what one's credences and preferences are (another theme from KAIL).

Someone might complain that verbally articulated practical reasoning is a grossly over-intellectualized model of real-life decision-making. That may be so, but the point does not help standard decision theory, which is *intended* as a theory of decision-making by an ideally rational agent. Human decision-making at best approximates to that supposed ideal. Presumably, an ideally rational agent *can* articulate the reasoning by which it made its decision. We are therefore within our rights to ask what that reasoning would look like when the ideally rational agent puts it into words. If we are merely given the usual decision-theoretic calculation concerning the agent's numerically quantified credences and preferences, we face the same problem as before: the agent is making its decision by reasoning about its own subjective psychological states. As already observed, such reasoning typically misses the point. The fact that one is in such-and-such subjective states is a very bad reason for rescuing the child, even by the low standard of ordinary human performance. An agent who can articulate its practical reasoning by rehearsing the usual subjectively interpreted decision-theoretic calculation is seriously irrational, not ideally rational.

That argument may seem to prove too much. Surely comparing options in terms of their expected utilities is at least often the most rational way of making a decision,

---

[2] *Logical* validity also depends on the *form* of the premise and conclusion sentences. For example, '$a \neq a$' is logically inconsistent while '$a \neq b$' is not, where 'a' and 'b' are distinct but synonymous names. In this paper, for the sake of simplicity, I mostly gloss over tricky issues about the relation between contents and the sentences that express them in a given context, and their consequences for validity, because they do not make too much difference to the overall picture (Williamson 2021 discusses some of the complexities). However, I will be explicit about guises or modes of presentation of ways of doing something, because they are important for intellectualism about knowledge-how.

and remains manifestly cogent when articulated by the decision-maker. Indeed, but in such cases the premises and conclusion of the reasoning concern probabilities *on the available evidence* and utilities *on some relevant scale of the good*, which are far from equivalent to the agent's subjective probabilities and utilities. One indicator of this is that a normal agent can make sense of the supposition that she assumed the *wrong* probabilities and utilities in her calculation, not through lack of self-knowledge but through thinking about the world beyond her present subjective probabilities and utilities in unrealistic ways. Thus calculations like those of standard decision theory can play a significant role in articulating reasons for action, but only when they are de-subjectivized. What is most chilling about 'rational' decision theory on its standard subjective interpretation is not its use of numbers (typical of idealized model-building) but its failure to provide a kind of reasoning available to rational agents themselves.

Once the practical reasoning has been de-subjectivized like that, any grading of the probabilities and utilities at issue has been transferred into the *contents* of the premises and conclusion from the agent's *attitudes* to those contents. The sentences expressing the premises and conclusion of the argument express ordinary true or false propositions, not special graded entities, though some of the propositions may be *about* graded non-subjective probabilities and utilities. Correspondingly, the agent's attitude to the premises and conclusion can be *outright* belief, rather than just some graded credence (though the latter is not excluded). The beliefs discussed in this paper are *outright* beliefs.[3] We can continue to sum up the agent's practical reasoning as a verbally articulated argument from premises to a conclusion, while keeping in mind that this is typically a rational reconstruction, an idealized schematization of a messy and unarticulated psychological process. At least it is a rational reconstruction that the agent herself could in principle reasonably endorse, unlike the travesty offered by standard decision on its subjective interpretation.

A belief-based account of desire can supply just such reasoning that the agent herself could reasonably endorse. If, for some Φ, the desire that P consists in the belief that $\Phi(P)$, then '$\Phi(P)$' may figure as the missing premise, to combine with the other premise 'If I do A, I'll bring it about that P' to obtain the conclusion in an ordinarily valid way. Typically, the argument will not be *deductively* valid, but very little ordinary reasoning is deductively valid. A lower level of defeasible truth-preservation from premises to conclusion will do.

Of course, to apply a standard of deductive or non-deductive truth-preservation to practical reasoning, truth-evaluable premises do not suffice; we also need a truth-evaluable conclusion. It must be closely related to 'I do A'. However, if '$\Phi(P)$' is not 'I desire that P' but something more impersonal, like 'It would be good that P', then what '$\Phi(P)$' and 'If I do A, I'll bring it about that P' defeasibly imply is less the pre-

---

[3] David Lewis's objections to a belief-based account of desire were developed within a graded framework; see Lewis 1988 and 1996, and for criticism Byrne and Hájek 1997 and Bradley and List 2009. For more on outright versus graded conceptions of belief see Williamson 202X. For an attempt to fit knowledge into a framework which tailors content to graded belief see Moss 2018; Pavese 2020b applies Moss's account to knowledge in action (that is not the same issue as the gradability of knowledge-how, for which see Pavese 2017). For reasons explained in my 202X, I am sceptical of the need for any such revisionary account of content.

diction 'I shall do A' than a more normative (but perhaps practical rather than moral) conclusion such as 'A is the thing for me to do' or 'I should do A'.[4] For convenience, I will use the 'should' formulation, with a suitable reading of 'should'. Since the agent has reasoned from premises she believes to that conclusion, she will also believe 'I should do A'. From that belief, it is a short—though not inevitable—practical step to forming the intention to do A.

The account does not require a uniform interpretation of 'Φ'. Even if we read 'Φ(P)' as 'It would be good that P', 'good' may have to take various senses in different contexts, sometimes moral, sometimes practical, and sometimes social, sometimes individual. In all these cases, one can read 'It would be good that P' as elliptical for the counterfactual conditional 'If P, it would be good that P', parsed as 'would(if P, it is good that P)', where 'would' is a local necessity operator, restricted to contextually relevant possible worlds.[5] The contextual variation in 'would' is another respect in which 'Φ(P)' is non-uniform in content. The counterfactual conditional is needed because 'it is good that' is a factive operator: 'it is good that P' is true only if 'P' is true.

Despite its tendency to sound philosophical in the abstract, 'good' is not to be understood as a highly theoretical term. As soon as a creature has to weigh different goods against each other—food, drink, sex, shelter, safety—a common scale is needed to do it on. It takes no great sophistication to regard some options as *better* than others. Still, 'would be good' sounds too measured for capricious or perverse acts when the agent's only answer to the question 'What did you do that for?' is 'I just wanted to'. Even smashing the crockery involves some low-level means-end reasoning. In such a case, the premise may be 'Smashing the crockery is the thing to do' rather than 'It would be good that the crockery is smashed' or 'It would be good that I smash the crockery'. Thus the desire to smash the crockery may consist in the belief that smashing the crockery is the thing to do.[6]

In such examples, the desire that P consists in a belief that disposes the agent to intend to bring it about that P. Of course, not all desires are very practical. Someone who does not know whether his ancestors were slave-traders may strongly desire that they were not without being tempted to intend to bring it about that they were not; backward causation is not an option—though he may wish it were. Even in such cases, there may still be a residual disposition to intend to bring something about, just thoroughly inhibited by the knowledge that the intention would be futile. If he idly imagines that backward causation *is* an option, but still, under that supposition, feels not the slightest urge to bring it about that his ancestors were not slave-traders, we may wonder whether he really has the desire.

On the envisaged account, which belief constitutes a given desire is contingent and changeable: it depends on the cognitive and conative economy of the agent at the

---

[4] David Velleman treats similar infinitives such as 'to be brought about' as merely expressing the corresponding attitude's direction of fit (1992: 16–17). He may stipulate such a meaning for his own use of such phrases, but in ordinary English they contribute to the truth-conditions of sentences in the usual way, as in 'She brings about whatever is to be brought about'.

[5] See Williamson 202Y and 2020 for more on 'good' and 'would' respectively.

[6] Velleman 1992 treats such examples as fatal to the assimilation of desire to belief.

time. Moreover, many different beliefs may constitute the same desire of the same agent at the same time. Thus the account does not strictly *equate* desires with beliefs: the relation is constitution, not identity.

A potential challenge to this comparatively loose connection between desires and beliefs comes from semantic arguments for desire as belief, which may require a tighter connection. In particular, Lloyd Humberstone has argued that 'certain sorts of propositional attitude ascriptions involving belief and desire cannot be expressed' in a language with just the belief operator 'B' and the desire operator 'W' (for 'want'), unless it is supplemented with an operator 'D' (for the placeholder 'desirable', his analogue of 'Φ'), 'with 'BDα' and 'Wα' equivalent (thus rendering 'W' itself definable)' (Humberstone 1987: 51). By contrast, the present account does not make 'desire' definable in terms of 'believe' and 'Φ'.

One of Humberstone's examples is (1):

(1) John thinks that Mary owns a mad dog, and he wants it to bite her.

Humberstone shows that attempts to formalize (1) fail if one uses just 'B', 'W' and quantifiers, but that (2) will do the trick once 'D' is available (op. cit. 53−6):

(2) B(∃x)(x is a mad dog owned by Mary & D(x bites Mary))

But (2) does not result from substituting 'BD' for 'W' ('want') anywhere, as a strict definition requires, since that would not yield the separation of 'B' from 'D' by the intervening material in (2). The same underlying issue arises if one speaker says 'John thinks that Mary owns a mad dog' and another comments 'He wants it to bite her'. The problem with (1) is not purely syntactic. One can felicitously replace 'thinks' in (1) by 'knows' or 'is sure', but not by 'doubts'. One can also felicitously replace 'wants' in (1) by 'prays for'. We seem to treat the first conjunct of (1) as licensing us to talk under the supposition that John's thought involves reference to an object, on which we can make our use of 'it' anaphoric. If so, (1) does not require a strict definition of 'W' in terms of 'B' and 'D'.

Humberstone also presents two quantifier-free examples, to which he applies the same technique as in (2), inserting material between 'B' and 'D'. The details are too intricate to present here, but in neither case does his treatment seem to require the sort of strict definition that would make trouble for the present belief-based account of desire.

On the present account, the agent is typically defeasibly disposed to infer 'I should bring it about that P' from 'Φ(P)', though the relevant reading of 'should' may vary. For a giant panda, 'Φ(P)' might be 'If P, I get to eat lots of bamboo shoots'. Given facts about giant pandas' digestive system, and a reading of 'should' sensitive to their needs, the move from the premise 'If P, I get to eat lots of bamboo shoots' to the conclusion 'I should bring it about that P' may well tend to preserve truth in a giant panda context. Equally, for a human, 'Φ(P)' might even be 'it would be for the general good that P'. Given a reading of 'should' sensitive to humans' collective needs, the move from the premise 'It would be for the general good that P' to 'I should bring it about that P' may well tend to preserve truth in a human context.

Some accounts of desires as belief-based have been argued to face counterexamples from perverse desires (Velleman 1992). But the present account permits perverse desires. Imagine Punk: generally (though defeasibly), believing that it would be idiotic to bring it about that P disposes him to intend to bring it about that P. No enthusiast for practical reasoning, he may even omit the intermediate step of believing that he *should* bring it about that P. Thus, since he reasonably believes that it would be idiotic to trash his flat, he is directly disposed to intend to trash his flat. Indeed, he succeeds: he intentionally trashes his flat. In such circumstances, his belief that it would be idiotic to trash his flat constitutes his desire to trash his flat. Thus the account allows Punk his perverse desire, as a result of his perverse hook-up between beliefs and intentions.

Punk has a perverse desire by combining a reasonable belief with an unreasonable hook-up between beliefs and intentions. One can also have a perverse desire by combining an unreasonable belief with a reasonable hook-up between beliefs and intentions. Imagine Saint: generally (though defeasibly), believing that it would be good for everyone that P disposes her to intend to bring it about that P. An enthusiast for practical reasoning, she always includes the intermediate step of believing that she *should* bring it about that P. Thus, since she unreasonably believes that it would be good for everyone that her parents experience intense suffering, she reasons that she should make her parents experience intense suffering, and so intends to make her parents experience intense suffering. Indeed, she succeeds: she intentionally makes her parents experience intense suffering. In such circumstances, her belief that it would be good for everyone that her parents experience intense suffering constitutes her desire that her parents experience intense suffering. Thus the account allows Saint her perverse desire, as a result of her original unreasonable belief.

Could one object that Saint's case violates the nature of belief? David Velleman writes (1992: 14):

> When someone believes a proposition, however, his acceptance of it is regulated in ways designed to promote acceptance of the truth: he comes to accept the proposition, for example, when evidence indicates it to be true, and he's disposed to continue accepting it until evidence indicates otherwise. Part of what makes someone's attitude toward a proposition an instance of belief rather than assumption or fantasy, then, is that it is regulated in accordance with epistemic principles rather than polemics, heuristics, or hedonics. An attitude's identity as a belief depends on its being regulated in a way designed to make it track the truth.

That is an extraordinarily idealized picture of belief. The world is full of people with intensely dogmatic beliefs (about who won an election or why vaccination is encouraged, for instance), which are impervious to new evidence. What makes them beliefs is that their holders are willing to *act on them*—sometimes, even when their lives are at stake. Saint's belief, against all the evidence, that it would be good for everyone that her parents experience intense suffering is no more dogmatic or unreasonable than beliefs held by millions of people today.

A normative variant of the objection is that beliefs, unlike desires, *should* be sensitive to new evidence. But some desires are commonly and properly described as 'irrational'; typically, an irrational desire is so *because* it is less sensitive than it should be to new evidence. For example, when I still want to eat the delicious-looking chocolate cake after learning that it is laced with arsenic, my desire is irrational. Although perverse desires are possible, they are often irrational.

Does irrationality take the same form for desires as for beliefs? Another line of objection to assimilations of desire to belief is that mutually inconsistent beliefs are problematic in a way in which mutually inconsistent desires are not. For example, two of your friends, X and Y, are in for the same job. You may well want X to get the job (it would mean so much to X!) and also want Y to get the job (it would mean so much to Y!), even though you know that they can't both get it. There is no irrationality in your attitude comparable to that of both believing that X will get the job (X is so good!) while also believing that Y will get it (Y is so good!). Suppose that your desire for X to get the job consists in your belief that if X got the job, it would be good that X got the job, and your desire for Y to get the job consists in your belief that if Y got the job, it would be good that Y got the job. Those two counterfactual conditionals are quite consistent. X getting the job and Y getting the job may both be good outcomes: even if the relevant standard of goodness puts no weight on your friendship with them, they may both be amongst the best candidates. Not only is the present account consistent with the rationality of desiring each of two mutually inconsistent outcomes, it explains *how* such desires can be rational.

Naturally, there is much more to be said about desires as belief-based, and the present account may turn out to need tweaking. Nevertheless, in this paper I will take it as a working hypothesis, to be judged by its fruits.

2 Four narratives

I will consider a sequence of four schematic narratives of practical deliberation and its outcome. The spirit of the first narrative is fully Davidsonian. The spirit of the final narrative is fully knowledge-first. The first narrative is transformed into the second by the belief-based view of desire just explained. The second narrative is transformed into the third by restricting the former to the case where all goes well on both the input and output sides. Conversely, that enables us to understand the second narrative in terms of the third as what remains from it when we drop the assumption that all goes well, in the knowledge-first spirit of understanding all cases in relation to the good case. The third narrative is transformed into the fourth by abstracting from the specification of means in the first three narratives and applying intellectualism about knowledge-how. Thus intellectualism plays a crucial role in reaching the fully knowledge-first and more general narrative of practical deliberation and its outcome.

Here is the first narrative. Imagine the schematic variable 'w' replaced by a specification of the means in practical terms, ready for you to execute.

<u>Narrative A   belief/desire/intention, means specified</u>.

Stage A1   You desire that P.
                You believe that w is a way for you to bring it about that P.

Stage A2   You desire to bring it about in way w that P.
                You believe that w is a way for you to bring it about that P.

Stage A3   You intend to bring it about in way w that P.

In this and subsequent narratives, you are assumed to have all the attitudes under the same *practical mode of presentation* of the way w, hooked up to your executive capacities, since that makes the attitudes most directly relevant to practical reasoning (Stanley and Williamson 2001). The diverse attitudes of believing, desiring, and intending can all be hooked up to the same practical mode of presentation, otherwise those attitudes could not properly interact as they do in practical reasoning; the mode of presentation presents the content, not the attitude to that content. Furthermore, we may assume for present purposes, if w is a way for you to bring it about that P, then it is possible for you to bring it about in way w that P, where 'possible' is understood as restricted to current practical possibilities, so that you *can* bring it about in way w that P.

Obviously, the transitions from stage A1 to stage A2 and from stage A2 to stage A3 are by no means deductive. The transition from A1 to A2 can fail when you believe that there are better ways than w for you to bring it about that P, or you have a stronger desire for something incompatible with your bringing it about that P. Conflicting desires may also block the transition from A2 to A3. Nevertheless, the transitions from A1 to A2 and from A2 to A3 plausibly work by default, when nothing intervenes to block them. In particular, the provision of what you believe to be a means (under a practical mode of presentation) eases the transition from the desire in A2 to the intention with the same content in A3.

Narrative A is fully in the spirit of belief-desire psychology, with intention as a further mental state. We could add extra conditions to defeat defeaters of the transitions between stages, but that would complicate the picture without achieving indefeasibility, since transitions from desires and beliefs to intentions are always subject to distractions, performance errors, and the odd meteorite, so never deductive. Keeping the picture simple and perspicuous is more conducive to insight.

To transform Narrative A into Narrative B, we apply the belief-based account of desire in Sect. 1. For concreteness, I take the relevant desires that P to consist in beliefs that it would be good that P, rather than beliefs with some other appropriate content. Not much hangs on that choice for present purposes, but it makes the transitions between stages easier to understand. I also insert an intermediate stage B2.5 between stages B2 and B3 to illustrate the role of beliefs about what one should do, already mentioned in Sect. 1.

Narrative B　　belief/intention, means specified.

Stage B1　　You believe that it would be good that P.
　　　　　　　　You believe that w is a way for you to bring it about that P.

Stage B2　　You believe that it would be good that you bring it about in way w that P.
　　　　　　　　You believe that w is a way for you to bring it about that P.

Stage B2.5　　You believe that you should bring it about in way w that P.

Stage B3　　You intend to bring it about in way w that P.

The transitions between stages are no more deductive than before, but still plausible defaults. Strikingly, the transitions from B1 to B2 and from B2 to B2.5 are both transitions wholly amongst beliefs, while the transition from B2.5 to B3 is simply from a belief that you should do something to the intention to do it (under the same practical mode of presentation).

We can articulate the agent's reasoning in the transition from B1 to B2 as the agent might put it in first-personal terms, thus:

Agent's first argument.

　　　It would be good that P.
　　　w is a way for me to bring it about that P.
So:　　It would be good that I bring it about in way w that P.

That looks like respectable default reasoning. The availability of this first-personal argument to the agent helps make the transition between stages A1 and A2 in the original narrative intelligible, because it clarifies for us as theorists the agent's rationality in making the transition. That reinforces the point made in Sect. 1, that a belief-based account of desire has an explanatory advantage over accounts that make belief and desire incommensurable attitudes.

Similarly, we can articulate the agent's reasoning in the transition from B2 to B2.5 as the agent might put it in first-personal terms thus:

Agent's second argument.

　　　It would be good that I bring it about in way w that P.
　　　w is a way for me to bring it about that P.
So:　　I should bring it about in way w that P.

That too looks like respectable default reasoning. The conclusion of the previous argument is the first premise of the new argument. The second premise of the new argument just repeats the second premise in the former; it offers reassurance on the feasibility of what the first premise says would be good. Again, the availability of this first-personal argument to the agent helps make the transition between stages A2

and A3 in the original narrative intelligible, because it clarifies for us as theorists the agent's rationality in making the transition (though here we also need the transition from B2.5 to B3). That adds to the explanatory advantage of a belief-based account of desire.

How much of a transition is there from B2.5 to B3? Is your intention to bring it about in way w that P really anything more than your belief that you should bring it about in way w that P (when those attitudes are under the same practical mode of presentation of w)?

We can test the potential gap by considering cases where several symmetrical ways w are equal best. The classic case is Buridan's ass, who has to decide which of two equally good bales of hay to eat first. Let 'P' be 'I satisfy my hunger'; one way is by eating the right bale first, the other by eating the left bale first. The ass can unproblematically make *both* transitions from B1 to B2. It truly believes both 'It would be good that I bring it about by eating the right bale first that I satisfy my hunger' and 'It would be good that I bring it about by eating the left bale first that I satisfy my hunger'. The ass is not irrational in both desiring to satisfy its hunger by eating the right bale first *and* desiring to satisfy its hunger by eating the left bale first; analogous cases of conflicting rational desires were discussed in section 1. But the ass *is* in trouble if it both *intends* to satisfy its hunger by eating the right bale first and *intends* to satisfy its hunger by eating the left bale first. The two intentions would trip over each other at the point of action. On which side of this contrast does stage B2.5 fall?

The two potential beliefs at stage B2.5 are 'I should satisfy my hunger by eating the right bale first' and 'I should satisfy my hunger by eating the left bale first'. If such beliefs already amounted to the corresponding intentions, the beliefs would already be in conflict. In the circumstances, the default reasoning in the second argument displayed above (concluding 'I should bring it about in way w that P') is defeated, and *neither* conclusion is adequately supported by the premises. But the ass may quite rationally form either intention, though not both. Thus, if it is rational in its beliefs, it will form an intention without having reached the corresponding belief by the argument proposed for the transition from B2 to B2.5. If the ass can get itself to plump for one way rather than the other, say for the right bale, it may *then* be able to describe its new state in the words 'I should satisfy my hunger by eating the right bale first; I should not satisfy my hunger by eating the left bale first', but that does not amount to discovering the right answer to the normative question it might originally have asked, 'What should I do?'. On any reading of B2.5 on which it is often reached from B2 by the reasoning displayed above, we cannot also understand B2.5 as already tantamount to B3.[7]

---

[7] For similar reasons, the future-tense belief 'I'll eat the right bale first' is not the conclusion of valid practical reasoning, on any standard of validity which requires the conditional probability of the conclusion on the premises (when defined) to be more than 50%. The point can be strengthened by considering choices between arbitrarily many equally good options. (For a defence of the strong cognitivist view that an intention to A is a belief that one will A based on practical reasoning see Marušić and Schwenkler 2018.) Since the ass cannot fairly be charged with *invalid* practical reasoning, its intention to eat the right bale first is not the conclusion of reasoning at all. However, for all that, its intention might still constitute the belief that it will eat the right bale first.

The next task is to transform narrative B into narrative C by turning the beliefs in B into knowledge (the beliefs came as they should) and the intention in B into intentional action (the intention went as it should). Unlike narratives A and B, narrative C is specific to the good case, where all goes well on both the input and the output sides. That is an application of the knowledge-first methodology. We understand narratives A and B retrospectively as weakenings of narrative C, to allow for things going badly on either the input or the output side. Narrative C reveals the point of all that practical deliberation: for knowledge of good ends and knowledge of means to those ends together to lead to the realization of those good ends (in a not typically moral sense of 'good').

Here is the result of the transformation:

Narrative C: knowledge/action, means specified.

Stage C1 You know that it would be good that P.
        You know that w is a way for you to bring it about that P.

Stage C2 You know that it would be good that you bring it about in way w that P.
        You know that w is a way for you to bring it about that P.

Stage C2.5 You know that you should bring it about in way w that P.

Stage C3 You intentionally bring it about in way w that P.

The first-person reasoning in the transitions from C1 to C2 and from C2 to C2.5 is exactly the same as in narrative B; the difference is that your attitude to the premises and conclusion is now knowledge rather than belief. Of course, since the reasoning is non-deductive, your knowing the conclusion does not follow from your knowing the premises by a standard closure principle for knowledge under competent deduction. Nevertheless, under favourable conditions, you *do* know the conclusion as well as the premises, and such conditions obtain in the case narrative C describes.

We could not have reached narrative C directly from narrative A, for without the intervention of the belief-based account of desire we have no reason to treat the knowledge that it would be good that P in C1 as the good case of the desire that P in A1, or the knowledge that it would be good that you bring it about in way w that P in C2 as the good case of the desire that you bring it about in way w that P in A2. Without the belief-based account of desire, we could only have reached a hybrid between narrative A and narrative C, with desire on the first line of each of the first two stages and knowledge on the second line. Thus belief would be raised to knowledge, and intention to action, but desire would be left behind as mere desire. That is a dissatisfyingly disunified picture: if both belief and intention have corresponding good cases, why not desire too? As with practical deliberation, the belief-based account of desire enables us to see a much more satisfying picture.

The point of the final transformation is to abstract from the specific way w in narratives A-C, to see the bigger picture. But this is also the point at which intellectualism about knowledge-how enters. On that view, you know *how* to do A just in case

for some way w, you know *that w* is a way for you to do A. For example, you know how to respond affirmatively in Serbian if you know that saying 'Da' is a way for you to respond affirmatively in Serbian. For a trickier case, imagine being tightly and firmly tied up; you can see the knots, and are competent with the usual ways of unty-ing them, but are currently unable to carry them out. In the present circumstances, you cannot untie the knots. On one reading, you know how to untie them; on another reading, you do not. Correspondingly, on one reading the usual ways are ways for you to untie the knots; on another reading, they are not. For purposes of your practi-cal reasoning about how to escape, the latter readings are more relevant, and will be adopted here. In general, the knowledge-that underlying knowledge-how is usually assumed to be under a practical mode of presentation of the way w, one which is ready for you to execute.[8] It is controversial whether that assumption is baked into the semantics of 'know how to' or is merely a conversational implicature, pragmati-cally generated in typical circumstances. For present purposes, we need not decide the semantic issue, since our interest is in connections that hold 'for the most part' or in the good case, not in exceptionless entailments.

Like the belief-based account of desire, intellectualism about knowledge-how is of course highly controversial (for example, Bengson and Moffett, 2011). I have defended it at length elsewhere, and will not do so again here (Stanley and William-son, 2001; see also Stanley, 2011). I will make just one point. Anti-intellectualists typically treat the string 'know how to' as if it were an idiom; they do not attempt to derive its meaning from the separate meanings of 'know', 'how', the infinitive, and how they are put together. But even if that string really is an idiom, it must also have an ordinary compositional ('literal') reading too; that compositional meaning can be shown to be broadly intellectualist, like the meanings of 'know why to', 'know where to', 'know where to', 'know whether to', and so on. Thus anti-intellectualism implic-itly predicts that 'know how to' sentences are systematically ambiguous in a specific way. There are linguistic tests for such ambiguity. Does the anti-intellectualist predic-tion pass those tests? Anti-intellectualists rarely acknowledge that obvious challenge, let alone make a serious attempt to meet it.[9] At any rate, I will take intellectualism about knowledge-how as a working hypothesis.

The task is to abstract from the specific way w in narrative C. To do that naturally, we generalize out the variable 'w' by existential quantification. When we do that to the occurrence of 'w' in the sentence on the second lines of stages C1 and C2, we obtain this sentence:

(3) For some way w, you know that w is a way for you to bring it about that P.

But (3) just is the intellectualist paraphrase of 'You know how to bring it about that P'. In particular, that knowledge was already assumed to be under a practical mode of presentation of the way w, as perhaps required for knowing how. Thus we have

---

[8]For detailed discussion of what practical modes of presentation might be see Pavese 2015, 2019, 2020a.

[9]For attempts to provide a compositional anti-intellectualist semantics for 'know how to' sentences see Roberts 2009 and Santorio 2016, discussed in Pavese 2021b.

'You know how to bring it about that P' as the second line of stages D1 and D2 of the final narrative.

The variable 'w' also occurs at three other places in narrative C. It occurs in the first line of stage C2. Existentially quantifying 'w' as in (3), we obtain (4):

(4) For some way w, you know that it would be good that you bring it about in way w that P.

Strictly speaking, (4) does not entail (5):

(5) You know that it would be good that you bring it about that P.

For instance, 'It will be good that you bring it about that you are holding the ace of spades by performing the card trick you have just learnt' does not entail 'It will be good that you bring it about that you are holding the ace of spades'. What matters may be your display of conjuring skills; just picking up the card from the table would achieve nothing. Moreover, you may know all these facts. Thus 'You know that it would be good that you bring it about that you are holding the ace of spades by performing the card trick you have just learnt' does not entail 'You know that it would be good that you bring it about that you are holding the ace of spades', even if knowledge is closed under entailment.

In the present case, we do have (6) as well as (4), by stage C1/D1:

(6) You know that it would be good that P.

But, strictly speaking, not even (4) and (6) together entail (5). Even if you know that it would be good that your friend gets the money he needs, it may not be good that you bring it about that he gets the money he needs—because his hurt pride at having to accept the money will poison your friendship—although it would be good that you bring it about that he gets the money he needs by making an anonymous donation that seems to come from a wealthy admirer. However, there is no need to *deduce* narrative D from narrative C. If narrative D *typically* or *usually* applies when narrative C applies, that already makes narrative D of present interest. Plausibly, (5) typically holds when (4) and (6) hold together. Thus we can use (5) in narrative D.

The variable 'w' also occurs at stages C2.5 and C3. When we existentially quantify those sentences on it we obtain (7) and (8):

(7) For some way w, you know that you should bring it about in way w that P.

(8) For some way w, you intentionally bring it about in way w that P.

Presumably, bringing it about in way w that P entails bringing it about that P. Thus, if both 'should' and 'know' are closed under entailment, (7) entails (9):

(9) You know that you should bring it about that P.

Similarly, if the intentionality of doing is closed under entailment, (8) entails (10):

(10) You intentionally bring it about that P.

Even if those entailments are not exceptionless, they come close: (9) typically holds when (7) holds and (10) typically holds when (8) holds. We can therefore use (9) and (10) in a typical narrative. We have thus obtained:

Narrative D: knowledge/action, means unspecified.

Stage D1   You know that it would be good that P.
           You know how to bring it about that P.

Stage D2   You know that it would be good that you bring it about that P.
           You know how to bring it about that P.

Stage D2.5   You know that you should bring it about that P.

Stage D3   You intentionally bring it about that P.

That is a plausible description of cases when all goes well in acting on what you know. For example: you know that it would be good that the light is on; you know how to bring it about that the light is on; you know that it would be good that you bring it about that the light is on; you know that you should bring it about that the light is on; you intentionally bring it about that the light is on—you walk over to the switch and flick on the light.

We can also work back from narrative D to narrative C, obtaining C as the result of filling in some details in D, by specifying the way w. It is C that supplies the premises of the agent's own first-personal reasoning, their first and second arguments displayed earlier. In particular, C articulates the role of narrative D's knowledge-how in supplying a key premise for the agent's reasoning about means. That articulation rests exactly on the intellectualism: knowledge-how can supply a premise to practical reasoning *because* it is knowledge that. Its ability to integrate knowledge-how with practical reasoning in the good case is a further argument in favour of intellectualism.

Obviously, most agents will not go through the first or second first-personal arguments above in so many words. Nevertheless, by integrating information from different sources to reach the point of appropriate action, the rational agent has to do something functionally similar. If the first-personal arguments were quite alien to the agent's cognitive processes, with the former not summarizing the latter in even the crudest fashion, that would undermine our attempt to understand the agent as rational (see Section. 1 on 'rational' decision theory). The paradigm of intelligent behaviour is acting on what one knows, which in effect is using what one knows as premises for practical reasoning, just as in the first-person arguments. As theorists, we can take that verbally articulated paradigm as a model to understand the behaviour even of non-verbal creatures.

In short: by integrating intellectualism about knowledge-how with the knowledge-first approach, we can better understand the role of knowledge-how in practical deliberation.

3 Going deeper

The integration of knowledge and action arguably goes much further than described in this paper, which has focused on practical deliberation. It may also concern the role of knowledge, especially knowledge-how, in the controlled execution of action, especially in skill (Pavese 2016, Stanley and Williamson 2017; see more generally Fridland and Pavese 2020). Both knowing what one is doing and knowing how to do it may also be essential to the intentional aspect of action (Anscombe 1957, Pavese 2021a). Furthermore, fully intentionally bringing it about that P may even be a practical way of coming to know that P. That would all be consonant with the knowledge-first picture of knowledge as central to intelligent life. The present paper has contented itself with clarifying some aspects of the role of knowledge-how in a knowledge-first view of practical deliberation.

References

Anscombe, G. E. M. (1957). *Intention*. Blackwell

Bengson, J., & Moffett, M. (Eds.). (2011). *Knowing How. Essays on Knowledge, Mind and Action*. Oxford University Press

Bradley, R., & List, C. (2009). Desire-as-belief revisited, *Analysis*, *69*, 31–37

Byrne, A., & Hájek, A. (1997). David Hume, David Lewis, and decision theory, *Mind 106,* 411–428

Davidson, D. (1980). *Essays on Actions and Events*. Clarendon Press

Fridland, E., & Pavese, C. (Eds.). (2020). *Routledge Handbook of Philosophy of Skill and Expertise*. Routledge

Frost, K. (2014). On the very idea of direction of fit. *Philosophical Review*, *123*, 429–484

Hawthorne, J., & Stanley, J. (2008). Knowledge and action. *Journal of Philosophy,105,* 571–590

Humberstone, I. L. (1987). Wanting as believing' *Canadian Journal of Philosophy*, *17*, 49–62

Humberstone, I. L. (1992). Direction of fit. *Mind*, *101*, 59–83

Lewis, D. (1988). Desire as belief, *Mind*, *97,* 323–332

Lewis, D. (1996). Desire as belief II. *Mind*, *105*, 303–313

Marušić, B., & Schwenkler, J. (2018). Intending is believing: a defense of strong cognitivism, *Analytic Philosophy*, *59,* 309–340

Moss, S. (2018). *Probabilistic Knowledge*. Oxford University Press

Pavese, C. (2015). Practical senses. *Philosophers' Imprint*, *15*, 1–25

Pavese, C. (2016). Skill in epistemology, Parts I and II. *Philosophy Compass*, *11*, 642-660

Pavese, C. (2017). Know-how and gradability. *Philosophical Review*, *126*, 345–383

Pavese, C. (2019). The psychological reality of practical representation, *Philosophical Psychology*, *32,* 785–822

Pavese, C. (2020a). Practical representation. in Fridland and Pavese, 2020, 226–244

Pavese, C. (2020b). Probabilistic knowledge in action. *Analysis*, *80*, 342–356

Pavese, C. (2021a). Knowledge, action, and defeasibility. In Jessica Brown and Mona Simion (eds.), *Reasons, Justification, and Defeaters*. Oxford University Press

Pavese, C. (2021b). Knowledge how.*The Stanford Encyclopedia of Philosophy Summer* (2021 Edition). Edward N. Zalta (ed.), https://plato.stanford.edu/archives/sum2021/entries/knowledge-how/

Roberts, C. (2009). Know-how: a compositional approach. In Erhard Hinrichs and John Nerbonne (eds.), *Theory and Evidence in Semantics* (pp. 183–213) CSLI Publications

Santorio, P. (2016). Nonfactual know-how and the boundaries of semantics. *Philosophical Review*, *125,* 35–82

Stanley, J. (2011). *Know How*. Oxford University Press

Stanley, J., & Williamson, T. (2001). Knowing how. *The Journal of Philosophy*, *98*, 411–444

Stanley, J., & Williamson, T. (2017). Skill, *Noûs*, *51,* 713–726

Stocker, M. (1979). Desiring the bad: an essay in moral psychology. *Journal of Philosophy*, *76*, 738–753

Tenenbaum, S. (2007). *Appearances of the Good: An Essay on the Nature of the Good.* Cambridge University Press

Velleman, D. (1992). The guise of the good, *Noûs*, *26,* 3–26

Williamson, T. (2000). *Knowledge and its Limits*. Oxford University Press

Williamson, T. (2017). Acting on knowledge. In J. Adam, E. Carter, Gordon, & B. Jarvis (Eds.), *Knowledge-First* (pp. 163–181). Oxford University Press

Williamson, T. (2018). Knowledge, action, and the factive turn. In V. Mitova (Ed.), *The Factive Turn* (pp. 125–141). Cambridge University Press

Williamson, T. (2020). *Suppose and Tell: The Semantics and Heuristics of Conditionals*. Oxford University Press

Williamson, T. (2021). Epistemological consequences of Frege puzzles, *Philosophical Topics*, *49*(2), 287-320

Williamson, T. (2021). Knowledge, credence, and strength of belief', in A.K. Flowerree, & Reed, B. (Eds.). *Expansive Epistemology: Norms, Action, and the Social World.* Routledge

Williamson, T. 202Y: Knowledge, credence, and strength of belief. In Paul Bloomfield and David Copp (eds.), *The Oxford Handbook of Moral Realism*. Oxford University Press