Do we really need a knowledge-based decision theory?

Davide Fassio[1,2] · Jie Gao[1,2]

Abstract

The paper investigates what type of motivation can be given for adopting a knowledge-based decision theory (hereafter, KBDT). KBDT seems to have several advantages over competing theories of rationality. It is commonly argued that this theory would naturally fit with the intuitive idea that being rational is doing what we take to be best given what we know, an idea often supported by appeal to ordinary folk appraisals. Moreover, KBDT seems to strike a perfect balance between the problematic extremes of subjectivist and objectivist decision theory. We argue that these alleged advantages do not stand up to a closer scrutiny: KBDT inherits the same kinds of problems as alternative decision theoretic frameworks but doesn't retain any of the respective advantages. Moreover, differently from other knowledge-action principles advanced in the literature, KBDT cannot fully explain the intuitive connections between knowledge and rational action. We conclude that the most serious challenge for knowledge-based decision theorists is to provide a substantive rationale for the adoption of such a view.

Keywords Knowledge-based decision theory · Knowledge · Decision Theory · Rationality

Often the outcomes of our actions depend on facts not fully accessible to us. Decision theory provides us with a series of rules that can guide our decisions in situations of incomplete information. The most popular of these rules is operative when the probability distributions of possible states of the world are available—so-called decisions under risk. This rule tells us to *take the action with the highest expected*

✉ Davide Fassio
davidefassio@gmail.com

Jie Gao
philobaikal@gmail.com

[1] Department of Philosophy, Zhejiang University, Yuhangtang Road 866, Xihu District, Hangzhou, Zhejiang, China

[2] Research Associate, Faculty of Humanities, University of Johannesburg, Johannesburg, South Africa

value, where the expected value of an action is the sum of the values of its possible outcomes weighed by the probability of each outcome. Suppose you are playing a fair roulette wheel in a casino. You can bet $1 on any of the 38 numbers. If you guess the right number, you will win $10, otherwise you will lose what you bet. Should you bet your dollar? In this case, the expected value of betting is $(\$10 \times 1/38) - \$1 \approx -0.74$, while the expected value of not betting is 0 (keeping your dollar). The rule tells you that you shouldn't bet.

Assume that the appropriate or rational decision one should make is one that maximizes expected value. An important question is what kind of probabilities should figure in the calculation of actions' expected values. Standard views hold that the probabilities in question are (rational) credences, which are often taken to equal a (rational) prior conditionalized on what one is (reasonably) certain of.[1] More recently, some philosophers have suggested that the probabilities relevant for rational decision are given by conditionalizing a suitable prior on *the knowledge the agent has*.[2] The relevant probability of a proposition p relative to an agent S and a time t would be given by $P(p|K)$, where $P(.)$ is a suitable prior and K is the total knowledge S has at t (Schulz, 2017, pp. 463–464). A consequence of this probabilistic framework is that all known propositions are assigned probability 1, and all propositions known to be false probability 0. If I know that I am wearing a blue shirt, the probability of this proposition conditional on all that I know, including the known proposition that I am wearing a blue shirt, is 1; conversely, the probability that I am not wearing a blue shirt is 0.[3] The resulting decision-theoretic framework recommends maximizing expected value on the basis of this knowledge-based notion of probability.[4]

---

[1] Orthodox subjective Bayesian decision theorists (e.g., Howson and Urbach, 1993; Jeffrey, 1965; Lewis, 1981) equate the relevant probability with one's actual credences (degrees of belief), others (e.g., Broome, 1995, ch.5; Comesana, 2020) with rational credences.

[2] For variants of this view see Hawthorne and Stanley (2008, pp. 577–580), Moss (2018a), Schulz (2017), Weatherson (2012, pp. 77–82). A precursor of this view is Levi (1980). For critical discussions of the view see Comesana (2020), Dutant (forthcoming), Greco (2013), and Kaplan (2009).

[3] The kind of probability involved in this version of decision theory is equivalent to the notion of evidential probability recently made popular by Williamson (2000). This notion is standardly classified as a kind of 'objective' (e.g., Moss, 2018a, §9.1; Williamson, 2000, p. 10), 'epistemic' (e.g., Peterson, 2009, §7.4) or 'primitive' (e.g., Weisberg, 2011, §2.3) probability, measuring "the objective degree of justification a body of evidence E lends to a proposition p" (Schulz, 2017, p. 463). This notion of probability is the most essential and distinctive feature of knowledge-based decision theory. This notion is distinct from the more orthodox subjective notions of probability as degrees of (rational) confidence in a proposition, insofar as (rational) certainty is neither necessary nor sufficient for knowledge. It is also distinct from more objective notions of probability such as frequencies, since actual or hypothetical long-term frequencies may not be known. We would like to thank two anonymous reviewers for pressing us to clarify these points.

[4] According to a more demanding version of the view, defended by Weatherson (2012) (see also Dutant forthcoming, fn.8.), not only are the relevant probabilities conditional on one's total knowledge, but also the structure of decision tables must match one's knowledge: for instance, a state should appear on the table if and only if one doesn't know that it doesn't obtain. This makes a difference if some states are not ruled out by what one knows but nonetheless have probability zero (if *e.g.* there can be infinitely many equally probable states not ruled out by one's knowledge). These differences are unsubstantial for our following discussions.

Let's call the latter view *knowledge-based decision theory* (hereafter, KBDT). Upholders of this view argue that it has several advantages over competing theories of rationality. First, KBDT would naturally fit with the intuitive idea that being rational is doing what we take to be best given what we know.[5] The latter idea is often supported by an appeal to ordinary folk appraisals: intuitively it seems appropriate to support the reasonability of our actions with claims about knowledge. For example, it seems appropriate to claim that it was irrational for Mary to spend all her money, since she didn't know that her contract would be renewed, or that it is irrational for Jack to sell his lottery ticket, since he doesn't know that his ticket will not win. Moreover, knowledge seems to play a fundamental and exclusive role in rationalizing attitudinal responses such as praise and blame, as well as attributions of responsibility and culpability.[6] Assuming that knowledge plays such a key role in rational deliberation and decision-making, it seems to naturally follow that knowledge should also play a central role in decision theory, our most popular and successful theory of rational decision.[7]

A second motivation for KBDT comes from comparisons to alternative theories.[8] If we are interested in a theory of rationality, the relevant notion of probability cannot be too objective. Suppose that today there is a high objective chance that it will rain. If we factor objective probabilities in our decision, it's clear that (other things being equal) taking an umbrella will be the choice that maximizes expected values. However, suppose that I have no idea that there is a high objective chance of rain. As a matter of fact, for all I know (e.g., given my knowledge of recent weather forecasts) it's very likely that it will not rain. Given the available information, it seems irrational to carry the umbrella with me. A decision theory maximizing expected value given objective probabilities cannot capture our intuitions about which decisions are rational in such circumstances. On the other extreme of the spectrum, subjective Bayesian decision theorists hold that rational decisions should be based on what we in fact believe or our actual credences. This type of view is equally problematic. The following example from Greco nicely shows why:

> Suppose I am completely convinced that I am the target of a Martian conspiracy. Acting on this belief, I purchase large quantities of tin foil, so as to fashion hats that will render me invulnerable to Martian mind-scrambling rays. Even if my beliefs are probabilistically coherent, there is a perfectly natural, intuitive sense of "irrational" in which these beliefs and the actions based on them are irrational (2013, p. 82).

---

[5] Dutant (forthcoming, §1), Hawthorne and Stanley (2008), Moss (2018a, §9.1), Schulz (2017, p. 463). For a defense of constitutive connections between knowledge and rational action see, e.g., Fantl and McGrath (2002), Hawthorne (2004), Hawthorne and Stanley (2008), Hyman (1999), Stanley (2005), Unger (1975, ch. 5), and Williamson (2005, p. 231). This idea is also defended in indirect ways, by first arguing for principles linking rationality to other concepts such as motivating reasons or evidence, and then arguing that these concepts are co-extensional or tightly related to knowledge.

[6] E.g., Blome-Tillman (2017), Littlejohn (2020), Moss (2018a, pp. 197–201).

[7] Cfr. Schulz: "If decisions should be based on what one knows, knowledge should play a fundamental role in decision theory" (2017, p. 463). See also Moss (2018a, 189–190) for similar considerations.

[8] For variants of this argument see Greco (2013, p. 82) and Weatherson (2012, §1.1).

As this example makes clear, rational decisions must rely on rational epistemic states. In a decision-theoretic framework, this translates into the idea that the kind of probability we should use when we maximize expected value must be sufficiently sensitive to evidential considerations. Probability given what one knows seems to be the ideal candidate for this role, striking a perfect balance between the problematic extremes of subjectivist and purely objectivist decision theories. KBDT could be seen as a plausible middle ground able to dodge the negative consequences of either view. Of course, the above considerations do not exclude that other notions different from knowledge, such as justified belief or certainty, can play the relevant rationalizing role. However, as Weatherson (2012, §1.1) remarks, the notion of knowledge seems to be the most natural and more applicable to real-life situations compared to the alternatives.

In this paper, we argue that these two sets of motivations for KBDT do not stand up to closer scrutiny. We show that, contrary to how it may look at first, KBDT is not a plausible compromise between subjectivist and objectivist extremes. On the contrary, we argue that it is seriously controversial whether the view can avoid the same kinds of problems vexing alternative theories.[9] We also argue that KBDT doesn't retain any of the alleged advantages commonly attributed to such theories. Moreover, we argue that KBDT fares worse than other approaches in providing a natural account of the intuitive connection between knowledge and rational action. KBDT can explain some ordinary folk appraisals based on knowledge, but in general it fails to vindicate the role that knowledge plays in rationalizing several types of actions, attitudinal responses, and attributions of responsibility and culpability. The upshot of our discussion is that the adoption of a KBDT stands in need of further motivation. While our arguments are not meant to provide a refutation of the view, we think they challenge knowledge-based decision theorists to provide new, more substantive rationales for the adoption of such a view.

This is the plan of the paper. In Sects. 1 and 2 we argue that KBDT doesn't obviously avoid the problems and doesn't retain the alleged advantages of, respectively, more subjectivist and objectivist decision theories. In Sect. 3 we compare KBDT to knowledge-action principles and show that only the latter can account for the intuitive connections between knowledge and rational action. In Sect. 4 we conclude.

Before proceeding further some important remarks are in order. First, as it may be clear from our previous discussion, our arguments are not directed against the idea that knowledge is tightly or essentially related to rational decision. Nor are we opposed to the idea that the notion of knowledge plays a central role in our ordinary

---

[9] Let us be clear on the specific scope of our claim. We are not going to argue that KBDT is affected by exactly the same problems as other views. This claim is far too strong. As Greco's example shows, there may be specific data and intuitive cases that are problematic for other views but that could be easily accommodated by KBDT. Our point is rather that for any general type of problem affecting other views, KBDT also is affected by a version of it. For instance, Greco's example shows that subjectivist decision theories sometimes recommend decisions that we consider intuitively irrational. In the next section we discuss other cases in which KBDT recommends decisions that we would consider equally irrational. Thanks to an anonymous reviewer for encouraging us to clarify the exact scope of our arguments.

folk appraisals of actions and other attitudes. Our critical focus is exclusively on providing accounts of such ideas in terms of a knowledge-based decision theory.

Second, while decision theory has been mainly conceived as a theory of rational decision, we do not exclude that some version of such theory can capture other dimensions of normative assessment.[10] This entitles us to compare the prospects of KBDT to those of more objectivist decision theories. While objectivist decision-theoretic frameworks cannot provide adequate accounts of human rationality, it is commonly thought that these frameworks can represent more objective standards of appropriateness. So understood, our thesis is not just that KBDT fares badly as a theory of human rationality. Rather, our claim is that it is unclear what role KBDT could have in theorizing about *any* kind of normative assessment. Our goal is to challenge the motivation of such a theory regardless of what type of assessment this theory is supposed to capture.

Third, in this paper we only assess motivations for KBDT which do not rely on the support of independently contentious views and theories. Examples of theory-laden defenses of KBDT appeal, for example, to the idea of a natural extension of the knowledge-first project from epistemology to rational theory (e.g., Dutant forthcoming), or to the support that such theory would provide to contentious views such as interest-relative invariantism (Weatherson, 2012) and probabilistic knowledge (Moss, 2018a, Ch. 9). These motivations for KBDT are rhetorically weak for at least two reasons. First, they presuppose the truth of contentious views. Being associated with such views may well be considered by many as reasons *against* endorsing this theory.[11] Second, as we will see in Sect. 3, it is far from obvious that a KBDT constitutes the best way for such views to implement a connection between knowledge and rational action.

Fourth, KBDT claims that probabilities in a decision theory should be conditional on the total knowledge an agent has at a specific time. This claim is supposed to be neutral about, and compatible with, any specific account of the nature of knowledge. Our arguments in the following sections do not rely on any specific understanding of knowledge, and are compatible with both internalist and externalist accounts of epistemic justification. The problems we put forward do not ultimately stem from features of knowledge, but from the implementation of a knowledge-based notion of probability in a decision theoretic framework.

---

[10] Such a hypothesis is voiced in Moss (2018a, §9.3). For recent uses of decision-theoretical frameworks modeling moral assessments see, for example, Zimmerman (2008).

[11] To consider a paradigmatic example of this ambivalence, Moss (2018a, §9.4), in the context of a defense of KBDT, considers the potential support to interest-relative invariantism provided by knowledge norms of action as a "significant objection" to such norms, "one of the most costly consequences of knowledge norms of action", rather than an argument in their support.

1 KBDT and subjective decision theories

As we anticipated above, a number of philosophers have argued that KBDT avoids problems affecting subjectivist variants of the theory. In this section we argue that it is controversial whether KBDT can avoid versions of these problems (Sect. 1.1). Moreover, we argue that KBDT doesn't retain any of the main alleged advantages commonly attributed to these theories (Sect. 1.2).

1.1 Common problems

In the introductory section we have considered one important problem for subjective decision theories, namely, that such views can sometimes recommend decisions that intuitively we would consider irrational. Greco's example well illustrates this point. In that example, the agent is completely convinced that she is the target of Martians. If that agent maximized expected utilities based on her actual beliefs, she would end up fashioning and wearing weird tin foils hats supposed to shield her from Martian mind-scrambling rays. We would all judge that agent irrational. The lesson we can draw from this type of case is that not any actual belief can be a suitable input in a decision table. Only rational beliefs seem to be good candidates to ground rational action.

If we assume that knowledge implies rational belief, KBDT straightforwardly avoids Greco's specific version of the problem. However, it is at the very least controversial whether KBDT can avoid other versions of the same problem. Sometimes KBDT seems also to recommend decisions that intuitively many of us would consider irrational. These circumstances can be illustrated by cases in which a subject is irrationally unresponsive to immediate and inequivocable information. Consider the following case, inspired by a similar one in Gibbons (2013, p. 179):

> *MISSING KEYS*
> Grace is at home searching frantically for her keys. She remembers that she left them somewhere in the dining room, but doesn't remember where. She looks in the direction of the dining table. The keys are in clear sight, staring her in the face on the otherwise empty table. Yet Grace surprisingly fails to revise any of her doxastic attitudes and continues searching for the keys in other places.

Since Grace doesn't form new beliefs, she also doesn't come to know anything new supporting the claim that the keys are on the table. All Grace knows about the keys' location is that she left them somewhere in the dining room, but this knowledge is compatible with the keys being anywhere in the room. So, the probability that the keys are on the table conditional on Grace's total knowledge is as low as that the keys are in any other place (on the cupboard, the windowsill, the sofa, the TV…).

Given the low knowledge-based probability that the keys are on the table, the choice that maximizes expected value is to continue her search. Thus, according to KBDT, Grace's decision should count as perfectly rational. However, to our ears (as well as Gibbons' and many others), this conclusion sounds highly counterintuitive, at least as unacceptable as that in Greco's Martian example. Intuitively the only rational thing for Grace to do would have been to stop searching and take the keys.[12]

Here someone may argue that an appeal to some form of external defeat could help KBDT to provide the intuitive verdict about the case. For instance, the fact that the keys are directly in Grace's visual field could defeat her knowledge that the keys could be anywhere in the dining room, or any relevant knowledge about the keys' location she may have had prior to her search. But if Grace no longer has the same knowledge, the relevant probabilities in KBDT may not be the same. This in turn would affect the evaluation of which action is rational according to this theory.[13]

Unfortunately an appeal to external defeat cannot help KBDT. First, let's grant that Grace's visual experience defeats some of her knowledge about the possible location of the keys. Still, it seems plausible that some relevant knowledge would survive defeat. For instance, it seems highly implausible that Grace's perceptual experience could defeat her memory that the day before she left the keys somewhere in that room. Even if this were the only bit of knowledge about the keys' location spared by the defeater, the probability that the keys are on the table conditional on that knowledge wouldn't be any higher than that the keys are in any other place in the room. So, the choice that maximizes knowledge-based expected value would still be to continue her search.

Consider then an even more radical hypothesis: Grace's perceptual experience completely defeats every bit of knowledge concerning the location of the keys. Far from solving KBDT's problem, this hypothesis could make things even worse for the view. This is because, even in that case, the probability that the keys are on the table conditional on Grace's total knowledge couldn't get any higher, and would even risk being lower. Principles of Regularity and Indifference (or, alternatively, Maximum Entropy), in combination with a lack of relevant information, would determine a regular distribution of probabilities evenly spread across all the possible places the keys could be, thereby further lowering the chances that the keys are on the table. In sum, since Grace didn't have any initial information indicating that the keys are on the table, no amount of defeated knowledge could modify probabilities in such a way as to make it more probable that the keys are on the table than in any other place. Less knowledge can only make the probability that the keys are on the table lower, and KBDT would still oddly recommend Grace to continue her search.

A moral that has been drawn from cases such as MISSING KEYS is that epistemic grounds that make a decision rational do not exclusively supervene on knowledge. As argued by several authors, rationality can be partially affected by features external to our doxastic states (and thus our knowledge), such as perceptual

---

[12] Other examples in the literature can be used to make the same point. See, e.g., Gibbons (2006, p. 22), Lord (2018, pp. 72–73), Zimmerman (2008, p. 36).

[13] Thanks to an anonymous reviewer and Masashi Kasaki for pressing us to address this worry.

experiences, easily available information, or facts one should have known.[14] If so, KBDT couldn't avoid the same kind of problem affecting subjectivist decision theories. This view would fare slightly better than subjective decision theory on this score, but couldn't avoid delivering similarly counterintuitive verdicts about which decisions count as rational in a restricted range of cases.

However, we should add a caveat. Like other authors, we share the intuitive judgment that in MISSING KEYS and similar cases discussed in the literature the agent's decision is irrational. But we are aware that philosophers have diverging intuitions about such cases. Those who disagree with our intuitive judgments and find Grace's reaction fully rational will not find it equally obvious that the kind of problem affecting subjectivist theories also affects KBDT. So, we must admit that our current point has only a limited argumentative force. Nevertheless, the current disagreement about the relevant intuitive judgments and the related disagreement about whether rationality is affected by non-doxastic conditions suggest that it is at least controversial whether agents in this sort of cases are fully rational, and thus whether KBDT can avoid versions of the problem.

One may suggest that a modified version of KBDT could avoid the problem. For example, we could consider a view according to which probabilities relevant for rational decision are conditional on facts that an agent is in a position to know. In the above examples, while Grace doesn't know facts that would support suspending her search, she is in a position to know such facts (e.g., that the keys are on the table). This seems a legitimate strategy to avoid the above problem, and indeed one favored by authors discussing this sort of example.[15] While it is not the primary concern of this paper to consider alternatives to KBDT, we would like to point to a crucial problem for the implementation of the above proposal into a decision-theoretic framework.

Several philosophers have recently pointed out that notions of potential accessibility such as 'being in a position to know' do not allow conjunction agglomeration.[16] An agent can be in a position to know a thing and in a position to know another thing, but from this it doesn't follow that that agent is in a position to know their conjunction. Standard examples involve agents who are in a position to know p and in a position to know that they don't know p, but not in a position to know that [p and they don't know p]. A consequence of this agglomeration failure is that sets of facts that an agent can be in a position to know can belong to separate partitions of an agent's potential perspective which support incompatible probability distributions. For example, an agent may be in a position to know a set of facts F^1 that make it 90% probable that p and in a position to know a set of facts F^2 that make it 10% probable that p, but not in a position to know the conjunction of F^1 and F^2. As a result, her potential perspective is incompatible with the attribution of unique

---

[14] E.g., Gibbons (2006, 2013), Goldberg (2018), Harman (1986), Littlejohn (2009), Lord (2018), Zimmerman (2008).

[15] Gibbons (2006, p. 22; 2013, pp. 179–180); Lord (2018, §3.3).

[16] Heylen (2016), Rosenkranz (2016, 2018), Williamson (2000, p. 203). See also Skorupski (2010, §2) for an earlier discussion of the problem.

coherent probability distributions to the possible states *p* and not-*p*. This makes the implementation of the present proposal within a decision-theoretic framework seriously problematic.[17]

1.2 Advantages of subjective decision theories that KBDT lacks

It is controversial whether KBDT can avoid problems affecting subjective decision theories. Can it at least retain some of their advantages? We argue that it doesn't. Below we consider the two main advantages commonly attributed to subjectivist decision theories and argue that KBDT doesn't retain any of them.

1.2.1 Guidance

Several authors share the thought that a good normative theory shouldn't merely provide an explanation of what makes an action right or wrong. Such a theory should also provide a decision procedure that may reliably guide agents to appropriate decisions. Lack of guidance is usually considered a serious pitfall in a normative theory.[18] A standard motivation to prefer subjective decision theories to more objective ones has precisely to do with considerations about guidance. Objective decision theories enjoin us to maximize value on the basis of facts we may have no idea about. Given our inevitable ignorance and uncertainty about facts of the world, we often can't be guided by such theories to make right choices. In many circumstances doing the right thing would be just a matter of luck. By contrast, it seems that a decision theory maximizing expected value given what is probable conditional on information accessible from the subject's standpoint could serve as a guide for one's decisions in all or most circumstances.[19]

Arguably, at least some versions of subjective decision theory can better serve as guides than a KBDT. This is particularly clear when we consider cases in which the verdicts of this theory dramatically fail to supervene on our internal mental states.

---

[17] Fassio (2021) further discusses the problematic consequences of this kind of agglomeration failure for this and other views in the literature. An anonymous reviewer suggests a possible fix to this problem. The suggestion is that probabilities relevant to rational decision-making should be conditional on the conjunction of all the facts that the agent via updating on their total evidence is in a position to know individually, even though the agent is not in a position to know them in conjunction. We agree that this view would avoid the problem. However, the normatively relevant total set of facts would not be accessible to the subject as a whole. This sort of decision theory would be far from satisfactory as an alternative to subjectivist versions of decision theory, since the relevant probabilities could be inaccessible and thereby could not guide agents that cannot agglomerate the relevant information. The attribution of normative relevance to facts that are beyond the ken of the agent would make such view more akin to forms of objectivist decision theory that we will discuss in Sect. 2.

[18] For an overview see Timmons (2012, § 1.3 and §1.7).

[19] These considerations are not restricted to decision theory. One of the main historical motivations for the loss of popularity of classic consequence utilitarianism and the corresponding rise of expected consequence utilitarianism has been precisely the need to obviate the lack of guidance of the former theory. Moreover, considerations about guidance have been recently used in support of perspectivist views about rationality and overall obligations. See, for example, Gibbons (2013) and Lord (2018, pp. 86–87).

Consider a specific example.[20] Suppose that you desperately want to eat a chocolate bar. You remember that earlier that day you saw one bar in the kitchen and another one in the living room. From your internal perspective there is an equal, very low, chance that any one of the two bars is not where you remember you saw it last time. In this moment you are closer to the living room, and since you are quite lazy, from your standpoint it seems slightly less costly for you to go and get the bar in the living room. Unfortunately things are not as they look: the bar in the kitchen is still over the fridge but, unbeknownst to you, your partner came home earlier than usual today and ate the one in the living room. In this scenario:

- you know that there is a bar in the kitchen.
- you don't know that there is a bar in the living room, though that proposition is very likely given what you know (say, given your knowledge the probability that your partner comes home earlier and eats the chocolate bar in the living room is 1%).

The following table represents the situation according to a KBDT—where K is the proposition that there is a bar in the kitchen and L is the proposition there is a bar in the living room, getting the bar has utility 101, not getting it has utility 0, walking to the kitchen has utility -1:

	K & L (Pr: 0.99)	K & not-L (Pr: 0.01)	Not-K & L (Pr: 0)	Not-K & not-L (Pr: 0)
Go to the kitchen (Gk)	100	100	0	0
Go to the living room (Gl)	101	0	101	0

According to KBDT, the expected utility of going to the kitchen outweighs that of going to the living room (eu(Gk):100 > eu(Gl):99.99). You should walk toward the kitchen. But it is clear that there is no way for you to be guided by KBDT to the right decision. From your internal perspective, the only rational thing to do is to go to the living room.[21]

Standard subjectivist decision theories can easily avoid this problem. By way of example, consider a standard Bayesian decision theory according to which the probabilities in a decision table should fit the degree of internal evidential support. In the above case, from your internal perspective there is an equal, low chance that any one of the two bars is not where you remember you saw it last time. So, according to this

---

[20] A more sophisticated version of the example has been recently introduced by Dutant (forthcoming, §2) in the context of a discussion of the New Evil Problem. Dutant also discusses an alternative theory that would avoid the problem. The theory is sufficiently different from KBDT to deserve a separate treatment, which unfortunately we cannot provide here. A similar example can also be found in Littlejohn (2019), who attributes the originality of such cases to Dutant. See also Moss (2018a, §9.2) for a different sort of example leading to the same conclusion.

[21] Let us stress that the present argument doesn't rely on any specific understanding of knowledge, and is compatible with any form of internalism and accessibilism about justification. The failure of guidance is not due to features of knowledge, but to the specific type of probability involved in a KBDT.

theory, K&not-L and not-K&L have the same low probability, and the decision that maximizes expected value is to go to the living room. Moreover, since probabilities are determined by your internal perspective, the theory can guide you to the proper conclusion.

More generally, since subjective decision theories calibrate probabilities on the internally accessible perspective of the agent, it is predictable that in most cases these theories can guide a rational agent to the choices that maximize expected value. In contrast, to the extent that the difference between knowledge and justified belief falling short of knowledge is not internally detectable, KBDT can imply stark asymmetries between knowledge-based probabilities and what internally appears to the agent. These asymmetries determine cases in which agents would be required to do the opposite of what their internal perspective suggests. The 'chocolate bar' case illustrates one such asymmetry: from the agent's internal perspective K&not-L and not-K&L have the same chance of obtaining but, given an internally undetectable difference in knowledge, in a KBDT the respective probabilities are different. If so, the theory is inevitably less guiding than subjectivist decision theories.

It is worth stressing that the type of situation illustrated by the above example is neither unimportant, nor isolated. We could find ourselves in similar situations in many cases in which our choices are partially based on rational but mistaken beliefs. More precisely, this type of situation could arise every time our choices depend on two propositions that we take to know, when actually only one is known.

Guidance objections are not unfamiliar to upholders of knowledge norms and knowledge-based decision theorists. In response, Williamson (2005, 2008) observes that the same sort of objection applies to any norm of action and decision, including subjective decision theories. As long as agents may be unaware, mistaken or uncertain about what they actually believe (and not just what they know), the local requirements of such norms may not be fully transparent to agents, who thereby may fail to be properly guided by such norms in specific circumstances. Following Williamson, several authors have argued that, given transparency failures, no theory of action can guarantee perfect guidance.[22] Schulz (2017, p. 479) observes that virtually any decision theory is going to face a problem of this kind.

While we do not have much theoretical stake in the present dispute, we think that there are various ways to resist such a radical skepticism about guidance. First, even if we admit that any theory may fail to guide agents in specific circumstances, it seems clear that some would fare better than others. It seems quite plausible that in normal circumstances an agent will find more helpful a theory that tells her to maximize value given what she thinks is the case rather than given the objective state of the world, or what she knows. In this respect, guidance failures of subjective

---

[22] Hawthorne and Srinivasan (2013), Smith (2010, pp. 84–86), Srinivasan (2015). For a discussion of the objection and reply in the context of a KBDT see Moss (2018a, p. 196) and Schulz (2017, p. 479). Besides arguing that the same objection applies to competing views, Moss also stresses that KBDT may better capture more objective kinds of assessment. We will discuss this possible response in the next section (Sect. 2).

decision theories are less frequent, limited to the uncommon cases in which we are unaware of internal attitudes relevant to our decisions. Moreover, we also think that the above objections cannot easily apply to a decision theory whose inputs are exclusively constituted by information that is immediately and introspectively accessible to the agent in a circumstance. As long as wakeful agents can introspectively access some seemings and thoughts, they can always use them to guide their decisions.[23]

1.2.2 Three-options cases

Another important motivation for perspectivist views in general and subjectivist decision theories in particular comes from the consideration of so-called three-options cases.[24] Consider the following example:

> *MINESHAFT*
> Ten miners are trapped either in shaft A or in shaft B, but we don't know which. Flood waters threaten to flood the shafts. We have enough sandbags to block one shaft, but not both. If we block one shaft, all the water will go into the other shaft, killing any miners inside it. If we block neither shaft, both shafts will fill halfway with water, and just one miner, the lowest in the shaft, will be killed (Kolodny & MacFarlane, 2010, p. 115).

Intuitively, in MINESHAFT the only reasonable and responsible thing to do is to block neither shaft. This seems true in spite of the fact that we know that this option is suboptimal from a purely objective standpoint. This indicates that our perspective plays an important role in determining what we should do—as perspectivists hold. This point is even more straightforward when framed within a decision theory. Assuming that in MINESHAFT we attribute equal value to each life, given our partial state of information blocking either shaft has expected value 5, while blocking neither shaft has expected value 9. While we know that there is one choice that will produce more actual value (save ten lives), intuitively what we should do is to maximize *expected* value.

One may expect that, like other forms of perspectivism, KBDT can also accommodate intuitions about three-options cases. Unfortunately this is not the case. While the adoption of some form of perspectivism seems to be the most promising strategy to account for our intuitive judgments in such cases, recently several philosophers have recognized that variants of such cases can be devised for at least some perspectivist views.[25] KBDT is one of these views. Consider a specific example:

---

[23] Of course the latter view would have other important problems. See Mason (2013, p. 7), Zimmerman (2008).

[24] Cases of this type have been discussed by, amongst others, Broome (2013, pp. 37–38), Gibbons (2013, p. 131), Jackson (1991, pp. 462–63), Kiesewetter (2011, 2017), Parfit (2011, pp. 159–160), Regan (1980, p. 265), Ross (2012), Zimmerman (2006, 2008, 2014). See in particular Jackson and Zimmerman for discussions of such cases within a decision-theoretic framework.

[25] See, e.g., Kiesewetter (2018, p. 109), Littlejohn (2009), Mason (2013, p. 7), Smith (2011).

MINESHAFT-II

All details of the case are as in the original mineshaft, except that in this scenario we know information that, if examined carefully, would make it highly probable in which shaft the ten miners are located (say, .99 probable that they are in shaft A). Moreover, a very reliable source has made us aware that we possess this relevant information. Unfortunately, due to contingent factors of the situation (time constraints, cognitive limitations…), we are not in a position to reach any significant estimate of where the miners are likely to be given what we know.[26] We must urgently take a decision, but at the moment we cannot tell in which shaft the miners are.

Conditional on our total knowledge, it is highly probable that the miners are in shaft A. According to KBDT, in this scenario the only rational choice (the one that maximizes knowledge-based expected utility) would be to block shaft A.[27] This verdict doesn't sound less counterintuitive than the one in the original MINESHAFT. It seems that given our contingent inability to reach any helpful estimate of where the miners are likely to be, the only reasonable and responsible choice for us is to block neither shaft.

Moreover, a popular line of argument shows that one could hardly coherently endorse a theory such as KBDT without being forced to make irresponsible decisions.[28] As long as we are aware that according to KBDT blocking neither shaft fails to maximize expected utility, if we were to endorse KBDT as the correct theory of rational action we would also believe that blocking neither shaft is impermissible. By taking this option, we would be doing something we think we shouldn't do. But doing what one believes one shouldn't do is a form of akratic behavior, commonly considered a paradigmatic instance of irrationality. Thus, by endorsing KBDT we would face the following uncomfortable situation: either take the only rationally permissible option—viz. block one of the shafts—which however seems also a deeply irresponsible one; or alternatively opt for the obviously responsible choice of blocking neither shaft, but at the cost of doing what we think is wrong, thus being akratic and overtly irrational.

It is a debated issue whether there is a view that can avoid problematic three-options cases, and what this view is. However, views that identify the relevant perspective with conditions that are more easily and transparently accessible to the agent seem to fare much better than KBDT. For instance, Zimmerman (2014, pp. 69–76) argues that a theory that limits the normatively relevant information to what

---

[26] These factors can include limits of time and cognitive and computational skills (Smith, 2011, p. 5 and fn.14; Williamson, 2000, p. 203), and lack of transparent access to the relevant information (Hawthorne & Srinivasan, 2013; Williamson, 2000, ch.4–5).

[27] Assuming that it is .99 probable on total knowledge that the ten miners are in shaft A and that each life is worth 1 utility, KBDT entails the following expected utilities and ranking of choices: eu(block A)=9.9 > eu(block neither shaft)=9 > eu(block B)=.1.

[28] Kiesewetter (2017) and Zimmerman (2008, 2014) discuss variants of this argument applying to moral and all things considered obligations. Notice also that this type of argument can be easily extended to knowledge-based alternatives to KBDT, such as the higher-order theory of Shultz (2017) and the indirect strategy of Dutant (forthcoming, §3).

an agent justifiably believes would avoid this kind of problem. A subjectivist decision theory whose inputs are restricted to introspectively accessible information would also do the work. By excluding all information not immediately and transparently accessible to the agent, we may ensure that the verdicts of a decision table fit our intuitive judgments about what choices count as rational and responsible in three-options cases.[29]

By way of example, consider a decision theory whose inputs are exclusively constituted by information that is immediately and introspectively accessible to the agent. Let's define as *immediately and introspectively accessible* any information that an agent has the physical and cognitive ability to retrieve and use to reach significant estimates about the situation at the moment of decision. It is easy to verify that no case such as Mineshaft-II can be devised for such a view. If due to contingent physical and cognitive factors of the situation we cannot retrieve and use some information to reach significant estimates of where the miners are likely to be, by definition that information is not immediately and introspectively accessible to us, and this theory rules it out as irrelevant to our rational decision. The theory will thus recommend us to block neither shaft. For the same reason, if an external source told us that we possess some useful information that we cannot retrieve and use, this information would not be relevant to our current rational decision.[30]

2 KBDT and objective decision theories

We saw in the previous section that KBDT doesn't fare well as a theory of what we should do in a subjective sense. Subjective notions of 'ought' and 'should' seem to play important roles in evaluations of reasonability and responsibility. However, these notions can also be used in other senses expressing more objective standards of appropriateness. Maybe KBDT could serve as a theory for one of these more objective standards.

The appeal to objective standards is one of the strategies adopted by upholders of knowledge norms in response to the challenge that such norms are not perfectly guiding. In particular, Sarah Moss claims that KBDT "may guide our assessments of the actions of others, as well as your retrospective assessments of your own actions, even in cases where it could not have usefully guided your actions directly"

---

[29] Could a version of KBDT restricting the relevant probabilistic base only to a subset of known facts to which one has immediate and transparent access at the moment of decision avoid the problem? While this view is not the primary target of our paper, we also think that it wouldn't be free from the problem. Mason (2013, pp. 7–8) argues that problematic versions of three-options cases can be devised also for views that consider as normatively relevant only a privileged subset of the actual attitudes of an agent, such as a subset of factive mental states, but exclude as normatively irrelevant other attitudes, such as rational but false beliefs.

[30] This is just an example. In general, subjectivist theories can avoid problems with three-options cases if they characterize the perspective relevant to rational decision in such a way as to exclude possible failures to access and use that perspective at the moment of decision. This subjectivist approach to three-options cases is well-known in the literature and is often recognized by both upholders and critics of subjectivist theories. See for example Mason (2013: p. 7), Zimmerman (2014: pp. 69–75).

([2018a](#), p. 196). The thought is that KBDT could serve functions traditionally attributed to objective senses of 'ought', such as explaining third-person assessments and informed retrospective self-assessments. For example, when we evaluate what someone should do in cases such as MINESHAFT, there seems to be at least one notion of 'ought' according to which it makes sense to say that the agent ought to block the shaft where all miners are located. This would definitely be the choice leading to the best actual result, saving the lives of all ten miners. Also, once the agent realizes where the miners were, say in mine A, there is one sense of 'should' in which it makes perfect sense for her to say 'I should have blocked shaft A'—we may call this 'regret' sense.

Objective standards can explain practices related to giving and receiving advice in similar ways.[31] In the MINESHAFT, an adviser fully informed about the situation would recommend that we block the shaft where the miners are. In so doing, the adviser would take an objective stance about the matter, completely ignoring the deliberator's subjective standpoint.

A further motivation for assuming more objectivist standards is that these tend to recommend choices leading to a higher degree of objective value. Such standards tend to factor most or only true propositions into a decision table. It may be argued that decisions taken on the basis of true propositions tend to be right more often than ones based also on false propositions.

While we are not opposed to the idea that there could be several more or less objective standards of appropriateness, the question here is whether KBDT can account for any of these standards. Granted that this theory cannot capture a subjective standard, could this theory at least express a more objective one, accounting for third-person assessments, retrospective self-assessments, advice or objective value maximization?

We don't think it can. It is pretty clear that our evaluations of other people do not track what an agent should do given what she knows. This standard would still count as too subjective. In the MINESHAFT, an external observer fully informed about the facts would probably excuse the deliberator for acting on her limited perspective and deciding to close neither shaft. Nonetheless, the observer would judge that the deliberator should block the shaft where the miners are, saving all their lives. Similarly, once the deliberator realized that all miners were in mine A she would properly say, in a regretful tone, 'I should have blocked shaft A'. These verdicts strikingly contrast with those of KBDT, according to which the deliberator should block neither shaft, the choice that maximizes expected value on the basis of what she knows at the moment of decision. In short, our assessments of others do not track their best decisions in light of their knowledge, and our retrospective assessments of our own actions do not track the best decisions given our knowledge at the time we made the choice.

The same point applies with even more intuitive force to the standards related to the practice of advice. As we anticipated above, if in the MINESHAFT our deliberator were asking for advice from an informed observer, the proper recommendation

---

[31] Graham ([2010](#), §2), Kolodny and MacFarlane ([2010](#), §2.2), Thomson ([1986](#), p. 179).

would be to block the shaft where the miners are. No reasonable person fully informed of the facts who knows where the miners are would ever advise acting according to what maximizes expected values given what the deliberator knows, thereby letting an innocent person die. Again, the objective stance taken by the advisor would not coincide with the verdict of KBDT. Recommending what maximizes expected value in light of the deliberator's knowledge at the time of the choice sounds as absurd as in light of any other deliberator's subjective standpoint. Thus, regarding the practice of advice, for KBDT what Judith Thomson said about subjectivism is valid:

> On those rare occasions on which someone conceives the idea of asking for my advice on a moral matter, I do not take my field work to be limited to a study of what he believes is the case: I take it to be incumbent on me to find out what is the case (Thomson, 1986, p. 179).

Kolodny and MacFarlane rightly observe that the same point Thomson makes about beliefs can be made for the agent's evidence (2010, p. 119), which on a natural construal of the expression simply means what one knows is the case.

Consider then the second type of motivation for objectivist standards: that such standards tend to determine a higher degree of objective value. Also on this score, KBDT seems not to fare particularly well. First, compare this view to a more objectivist standard, such as the following:

> (PDT) Choose option O if and only if O maximizes expected value in light of physical probability.[32]

Suppose you can decide whether to bet on the outcome of a coin toss. If the coin lands heads you win $10, otherwise you lose $20. For all you know, the coin is fair. This means that conditional on your knowledge the probability of each outcome is 50%. KBDT recommends not taking this bet (eu(bet) = −5 < eu(not bet) = 0). However, unbeknownst to you, the coin is double-headed. This means that the physical probability that the coin will land heads is 1. PDT recommends taking the bet (eu(bet) = 10 > eu(not bet) = 0). If we compare the two theories on the basis of which leads to a higher degree of objective value, PDT fares much better than KBDT. The MINESHAFT case provides another nice example. PDT attributes probability 1 to the true state of the world (e.g., that miners are in shaft A) and 0 to the other possible state. Thus PDT recommends an objectively better choice than KBDT: save all miners' lives rather than letting one die. The result can be easily generalized: if we assume that physical uncertainty is limited to cases that are relatively isolated or irrelevant in practice, in the long run acting according to PDT will have hugely better outcomes than acting on KBDT.[33]

---

[32] If physical determinism is right, this principle is equivalent to "Choose O if and only if O is the best option in light of all the facts (known and unknown)".

[33] If one thinks that our point could affect only a limited range of cases such as biased coin tosses and extreme moral scenarios, just consider the following case: The objective physical probabilities of a COVID-19 pandemic were already very high as early as November 2019. Earlier prevention measures would have avoided the spread of the virus, but unfortunately all knowledge available to the scientific

At this point one may argue that if we assess views according to how much objective value they bring about, KBDT is worse than purely objectivist views but at least it fares better than more subjectivist ones which factor in the decision table some false beliefs also. A possible rationale for this thought could be the idea that decisions taken on the basis of true propositions tend to systematically lead to better consequences than decisions also partially based on false propositions. But this is not obviously true. If we look exclusively at the long-term objective benefits, an agent who complies with KBDT doesn't necessarily fare any better than one conforming to more subjective versions of decision theory. If for example a fallacious form of reasoning leads to false beliefs that will lead to many more accurate attitudes which in turn will lead to higher objective value, a theory allowing that such false beliefs determine probabilities in a decision table will likely have more objective benefits than one disallowing such beliefs.

Indeed there are good reasons to think that in a wide number of situations conforming to KBDT brings about less objective value than some alternative more subjectivist theory. Several authors have argued that there are cognitive shortcomings that are practically beneficial. For example, Stich (1990) and Stephens (2001) argue that true beliefs are not always adaptive. Sullivan-Bissett (2017, §2) observes that while true beliefs are in general more likely to dispose a creature to act in ways which will satisfy its desires and wellbeing, there are cases of usually false beliefs which best serve our practical purposes, such as "beliefs produced via self-enhancement bias, partiality bias (different doxastic treatment of one's friends over strangers), and self-deception, to name just a few" (2017, p. 96). These beliefs serve important purposes for human wellbeing such as facilitating self-organisation, maintaining self-esteem and avoiding psychological damage. While these beliefs do not constitute knowledge, once implemented into our deliberative practices they tend to deliver the best practical results. Since KBDT excludes every false belief from being practically guiding, it also excludes this important subclass of false but highly helpful beliefs. As a consequence, the theory would provide a considerably reduced amount of objective value overall compared to the value we actually have.

In addition, there are cases in which it seems perfectly appropriate to factor into our decision propositions which we know are false. For example, Gao (2017) argues that in certain circumstances it can be perfectly reasonable (and sometimes even rationally required) for scientists to reason from theories known to be false, as long as these constitute good approximations to the truth for present purposes. For instance, in doing a calculation which does not require a high level of accuracy, a scientist who uses only known theories (e.g., quantum mechanics) as inputs of her

---

Footnote 33 (continued)

community (or anyone else) at the time was not nearly enough to provide even a slight probabilistic support for the catastrophic event. At that time the choices maximizing expected utilities on objective physical chances (e.g., early lockdown and massive testing) would have surely led to better results than the ones maximizing expected utilities on knowledge-based probabilities.

decision theory would spend much more time and energy, and hence be practically worse than one using a simpler but falsified theory (e.g. Newtonian mechanics).

Notice that we are not claiming that KBDT fares worse than any more subjectivist decision theory. We agree that KBDT would likely fare better in terms of objective value compared to a theory maximizing utility on any sort of unreasonable and false belief: if you believe that there are streets where in fact there are cliffs, your objective utility will likely be severely reduced. It is however controversial whether KBDT would actually fare better than a theory taking as practically guiding mostly false justified beliefs. If, for instance, advances in our understanding of the natural world were showing that most of our ordinary commonsense beliefs are false, these beliefs would have nonetheless proved to be appropriate from an evolutionary perspective, leading to human wellbeing and flourishing.[34]

Let's take stock of the conclusions in this and the previous section. In Sect. 1 we argued that KBDT doesn't fare well as a theory of what we should do in a *subjective sense*. In this section we argued that KBDT also doesn't fare well as a theory of what we should do in more *objective senses*. The view seems to be too objective to retain advantages of subjectivist theories (modest guidance, correct diagnoses in three-options cases), but also too subjective to retain advantages of objectivist theories (e.g., explanations of third-person and regret assessments, advice, objectively best choices). Since the advantages of subjectivism are disadvantages of objectivism and vice versa, KBDT also retains all the same kinds of problems of each theory. Far from being a plausible middle ground between the problematic extremes of subjectivist and purely objectivist decision theories, this view succeeds in catching all the respective problems while avoiding all the advantages.

3 Lotteries, blame and culpability

As anticipated in the introduction, a second important consideration advanced in support of KBDT is that this theory would provide a plausible account of the intuitive connection between knowledge and rational action. This connection is naturally suggested by a range of ordinary folk appraisals: it often seems appropriate to support or challenge the reasonability of certain actions and attitudes on the basis of claims about knowledge. These intuitive data have been traditionally deployed to defend principles such as the following[35]:

Knowledge-Action Principle (KAP) It is appropriate to treat the proposition that *p* as a reason for action only if you know that *p*.

---

[34] Evolutionary psychologist Donald D. Hoffman (2019) has recently argued that evolution does not favor accurate perception. He provides empirical evidence that not seeing accurately gives us a survival advantage, and therefore that our representation of reality evolved to hide reality from us, rather than accurately represent it. If Hoffman were right, a perspectivist theory of rational decision including mostly false justified beliefs would promote considerably more objective value than KBDT.

[35] For similar principles see Fantl and McGrath (2002), Hawthorne (2004), Hawthorne and Stanley (2008), Hyman (1999, p. 442), Stanley (2005), Unger (1975, Ch.5), and Williamson (2005, p. 231).

A further natural step leads to the conclusion that knowledge should also play a central role in our most popular and successful theory of rational decision, namely, decision theory. The step is very explicit in Schulz, who claims that "[i]f decisions should be based on what one knows, knowledge should play a fundamental role in decision theory" (2017, p. 463). Similar claims are endorsed by other proponents of KBDT such as Moss (2018a) and Dutant (forthcoming).

While apparently very natural, we do not think that the latter step is legitimate. This is because several ordinary folk appraisals allegedly supporting a normative connection between knowledge and action can be straightforwardly explained by principles such as KAP, but cannot receive any adequate explanation in terms of a KBDT. On the contrary, the latter theory seems to deliver the wrong verdicts about such cases. In what follows, we will focus in particular on three sets of intuitive data discussed in the contemporary literature: lotteries, negligence and blame, and attributions of responsibility and culpability.[36]

3.1 Lotteries

Naoko received a lottery ticket as a gift for her birthday.[37] She knows that the lottery is fair, that there are 10,000 tickets, and that the prize is $4,000. Intuitively in this situation it doesn't seem right to conclude that Naoko should sell her ticket for 50 cents based on the thought that the ticket will not win. Hawthorne and Stanley (2008, p. 579) argue that this type of case provides evidence in support of KAP: the best explanation for why it seems unreasonable to sell the ticket for 50 pennies seems to be that Naoko does not know that her ticket will lose. KAP provides a simple and straightforward explanation of the intuitive judgment about this and other similar cases.

One may expect that the same argument carries over KBDT. Unfortunately this is not the case. Assume that you know all the relevant facts about the case (the number of tickets, the price, that the lottery is fair). In a KBDT, Naoko's decision could be represented as follows[38]:

---

[36] These data are explicitly used by KBDT theorists as evidence for their view. We do not consider here a range of further data traditionally deployed in support of knowledge-action principles, such as direct criticisms of the reasonability of one's actions. An example is Hawthorne and Stanley's *restaurant* case, in which a driver is criticized for taking the wrong street on the basis that she didn't know that was the right direction. The reason is that even philosophers endorsing such principles tend not to consider these as compelling cases for KAP. For recent examples see Littlejohn (2020, §1) and Moss (2018a, p. 197), who claims that such cases are "suggestive but not decisive", since they are susceptible to alternative equally intuitive explanations (cfr. Brown, 2008, Cresto 2010, Fassio 2017, Gerken 2011, Neta 2009, Schiffer 2007).

[37] The present example is a variant of one discussed in Hawthorne and Stanley (2008, p. 572).

[38] For simplicity's sake in this representation of the situation we assume that in this specific situation utility is equal to monetary value. One may argue that this is an unrealistic assumption: monetary value has always a decreasing marginal utility due to risk aversion. This consideration may be correct, but doesn't affect the validity of our general point. We can obtain the same result by modifying specific details of the case, such as the number of tickets or the lottery price. Assuming decreasing marginal utility would just add unnecessary complications.

	The ticket will lose (Prob. 0.9999)	The ticket will win (Prob. 0.0001)
Sell the ticket	0.5	0.5
Keep the ticket	0	4000

In this case, the expected value of selling the ticket is higher than that of keeping it (0.5 > 0.4). Thus, in a KBDT, it is fully rational for Naoko to sell her ticket for few cents: this is the choice that maximizes expected value. This verdict conflicts with our ordinary folk appraisals about the case, according to which Naoko shouldn't sell her ticket for a few pennies. Far from providing the best explanation of such appraisals, KBDT provides diagnoses which conflict with such intuitive judgments.

A potential worry here concerns whether we should accept the verdict of our folk appraisals as correct. The verdict relies on the pretheoretical intuition that it is irrational to sell a lottery ticket for a few cents despite the unfavourable odds. This intuition doesn't seem to depend on how likely it is that Naoko will lose the lottery. As long as the lottery price is sufficiently high, it seems intuitively unreasonable to give up the chance of winning for a mere few cents. Only knowledge that the ticket is a loser could justify selling it. However, one may wonder, this verdict is not just incompatible with KBDT, but also with standard Bayesian decision theory, at least as long as an agent has rational beliefs about the circumstances. These may be grounds to challenge and explain away the ordinary intuitive judgments about such cases rather than rejecting KBDT and other decision theories.

While we share this worry, we also think that this doesn't affect our present point, which is to assess the motivations for adopting a KBDT. In particular, in this section we are considering the strength of the argument endorsed by upholders of KBDT according to which this theory, on a par with principles such as KAP, would provide a natural explanation of a series of ordinary folk appraisals. We just considered one of these appraisals, consisting in intuitions about lottery cases. The result is that, far from providing a good explanation of this type of appraisal, KBDT provides verdicts that are inconsistent with them. Of course, it may well be that these appraisals are irrational and fallacious.[39] But then, in this case also, the considered motivation for KBDT would be lost.

3.2 Negligence and Blame

A second type of ordinary appraisal that could be explained by principles such as KAP concerns how judgments of blame and negligence interact with knowledge. First, consider some examples in Hawthorne and Stanley:

> If a parent allows a child to play near a dog and does not know whether the dog would bite the child, and if a doctor uses a needle that he did not know to

---

[39] See Papineau (2019).

be safe, then they are prima facie negligent. Neither the parent nor doctor will get off the hook by pointing out that the dog did not in fact bite the child and the needle turned out to be safe, nor by pointing out that they were very confident that the dog/needle was safe (2008, pp. 572–573).

These (and similar) intuitive cases indicate that judgments of negligence are tightly related to knowledge. In particular, they indicate that in order to avoid charges of negligence, an agent should know the relevant facts (that the dog will not bite the child, that the needle is safe). Mere confidence is not enough to escape blame and criticism. These intuitions are straightforwardly explained by KAP: the fact that it is inappropriate to act on propositions you don't know provides a simple and straightforward explanation why it is inappropriate for parents to let their child play with a dog they *don't know* will not bite him, and for a doctor to use a needle she *doesn't know* to be safe.

Another, even more interesting and indicative example has been recently discussed by Lara Buchak:

> *CELL PHONE*
> You leave the seminar room to get a drink, and you come back to find that your iPhone has been stolen. There were only two people in the room, Jake and Barbara. You have no evidence about who stole the phone, and you don't know either party very well, but you know (let's say) that men are 10 times more likely to steal iPhones than women (2013, p. 292).

In this situation it seems rational to be more confident that Jake stole the phone. Nonetheless, it seems inappropriate to blame Jake for the deed. This seems to be true even if the probability that Jake stole the phone were higher (say, 99%). While it may be appropriate to perform certain acts on the basis of such statistical considerations (say, bet that the phone has been stolen by Jake), blame and resentment are not appropriate in this type of situations. Things would be different if you knew that Jake stole the phone—for example, if you saw him taking the phone. In that case, blame and resentment would be fully appropriate. These intuitions push us to accept the principle that we shouldn't blame people for ϕ-ing if we don't know that they ϕ-ed. Anything less than knowledge is not enough for blame, and nothing more than knowledge seems to be required. Once again, a variant of KAP ranging over attitudinal responses such as blame and resentment provides a straightforward explanation of the data.[40]

Assume that a decision theory could range over attitudes such as blame and resentment—or alternatively consider active expressions of blame and resentment

---

[40] For further discussions of the role of knowledge in assessments of blame see also Dietz (2018), Littlejohn (2020), Moss (2018a, 2018b). While here we focus exclusively on blame and resentment, the present point can be generalized to other emotions (e.g., shame, praise) and more in general to all reactive attitudes. Such attitudes and their attributions seem to be admissible in response to a certain fact only if one also knows that fact.

as options to be factored in the decision table.[41] As for lottery cases, the predictions of KAP and KBDT neatly diverge in the assessment of negligence and blame cases. Provided that the probability conditional on our knowledge that Jake stole the phone is sufficiently high, blaming and feeling resentment for Jake will maximize knowledge-based expected value. Thus KBDT will end up recommending these attitudes, in spite of the fact that you do not know that Jake stole the phone.[42]

The same point applies to attributions of negligence. KAP predicts that a doctor shouldn't use a needle if she doesn't know it is safe. But KBDT predicts that a doctor may use the needle even if she doesn't know it is safe. As long as the proposition that the needle is safe is sufficiently likely given what the doctor knows, using the needle will maximize knowledge-based expected value.

In sum, ordinary appraisals indicate that mere statistical grounds cannot justify attitudes such as blame and resentment and judgments of negligence. Knowledge that one performed the blameworthy or negligent action is also necessary. Against the verdict of ordinary folk appraisals, KBDT predicts that it could be permissible to blame someone and judge her negligent on the basis of mere statistical considerations. Again, far from providing the best explanation of such appraisals, KBDT provides diagnoses conflicting with such intuitive judgments. Once again, the best explanation arguments for KAP do not carry over KBDT.

3.3 Responsibility and culpability

A number of philosophers have recently argued that jurors in trials should not convict and punish a defendant unless they know that she is guilty of the crimes of which she is accused.[43] This claim is often supported by intuitive judgments about cases such as the following:

---

[41] A difficulty in assessing the rationality of 'reactive attitudes' such as blame, praise, gratitude and resentment is that these responses are not completely under our direct control. As Strawson (1962) observed, it may be practically impossible to control and stop reacting in such ways. These facts make it uncertain whether it is appropriate to input such attitudes as options in our decision theories. However we can bypass this issue by considering specific actions tightly related to such attitudes, such as explicit sincere expressions of blame.

[42] Moss has recently argued that in CELL PHONE "it is intuitive to say that you do not know that Jake probably stole your phone on the basis of facts about other members of his gender" (2018a, p. 198). We find Moss's claim very counterintuitive. We agree that from mere statistical grounds you cannot conclude that Jake stole the phone and blame him for doing so, but it seems to us pretty obvious that given what you know it is quite likely that Jake stole the phone. And it seems equally obvious that you could use this knowledge as a ground for other actions and attitudes, such as suspecting that Jake stole the phone and betting accordingly. Moreover, even if we grant Moss's point about knowledge of what is probable, we should keep in mind that probabilities factored in a knowledge-based decision theory are not known probabilities, but probabilities given what one knows. Moss's claim is compatible with the fact that conditional on your knowledge it is highly probable that Jake stole the phone.

[43] Variants of this view have been defended by e.g., Blome-Tillman (2017) and Littlejohn (2020). The view also aligns quite well with current legal practices.

PRISON YARD

100 prisoners are exercising in the prison yard. Suddenly 99 of them attack the guard, putting into action a plan that the 100th prisoner knew nothing about. The 100th prisoner played no role in the assault and could have done nothing to stop it. There is no further information that we can use to settle the question of any particular prisoner's involvement (Littlejohn, 2020, p. 5253).[44]

In PRISON YARD intuitively it is not permissible to convict and punish a prisoner chosen at random from the yard. One plausible explanation seems to be that the only evidence of guilt in this situation is statistical evidence. This type of evidence makes it highly likely that the prisoner is guilty, but is not enough to know that he is guilty. Only this knowledge would entitle attribution of responsibility and culpability to the prisoner and accountability for the crime committed.[45]

This explanation further strengthens the idea that knowledge and rational action are tightly interconnected. A knowledge-norm of action such as KAP receives further support from such data. Indeed, the norm that one shouldn't punish a defendant unless known guilty is a natural extension of KAP to the legal domain. However, this argument for KAP doesn't carry over KBDT. Suppose the jurors are aware of the facts described in PRISON YARD: in particular, they know that there is a high probability given what they know that a prisoner *P* chosen at random from the yard is guilty. Benjamin Franklin once said that it would be better that one hundred should walk free than one innocent person be sent to prison. Suppose that Franklin was right: if punishing a guilty person has value 1, then punishing an innocent has disvalue -100. Moreover, assume also that not punishing the guilty has a disvalue of -1. In a KBDT, the Jurors' decision can be represented by the following table:

	P is guilty (Prob. 0.99)	*P* is not guilty (Prob. 0.01)
Punish *P*	1	-100
Do not punish *P*	-1	0

Given the above value assignments, the choice that maximizes expected value is to punish *P* ($-0.01 > -0.99$). Contrary to both our intuitive judgments and our current legal practices, KBDT predicts that it is perfectly fine, and in some circumstances even obligatory, to convict and punish a defendant on mere statistical grounds, in spite of the fact that it is *not known* whether the defendant is guilty. All that KBDT requires to deliver punitive verdicts is that conditional on one's knowledge the defendant is very probably guilty.

---

[44] This and other similar cases have been widely discussed in philosophy of law. For an overview see Gardiner (2018).

[45] It is worth observing that the present considerations should not be confined to the legal domain. To borrow an example from Papineau (2019, §11), imagine that you notice certain changes in your spouse's behavior. You also learn from a reputable psychology magazine that it's highly likely that people behaving in this way are having an affair. While these statistical data may lead you to suspect your spouse of infidelity, they can't justify an explicit accusation of infidelity, no matter how high the relevant probability is.

The above value assignments are somewhat arbitrary, but this doesn't affect our general point. One may disagree with Franklin on the comparative disvalue of punishing an innocent, or with other features of this specific case. However, for any finite value assignment in the boxes, there are possible conditions under which KBDT predicts that one should punish a prisoner merely on the basis of statistical evidence. To get this result, we can just manipulate probabilities (e.g., by increasing the number of guilty prisoners in the yard) or add disvalue to letting the guilty prisoners go.

As for other types of cases considered above, KBDT prescribes decisions that neatly conflict with ordinary intuitive appraisals of the situations. Thus the theory is badly suited to provide the best explanation of these appraisals. Far from supporting KBDT, intuitive judgments about the conditions under which it is appropriate to hold a person responsible and culpable constitute an obstacle for the theory. While we can't exclude that upholders of KBDT could address this problem (for example, by explaining away the relevant intuitions), the crucial point of our discussion is that the theory fails to be motivated by an appeal to such ordinary folk appraisals.

3.4 A diagnosis of the problem and a reply to a possible objection

In this section we have considered a range of ordinary folk appraisals connecting knowledge to rational action. While these data seem to support principles such as KAP, upholders of KBDT conjectured that the same type of considerations may support the latter theory equally well. We have argued that this conjecture is false. While apparently very natural, the step from principles of appropriate deliberation to decision theory doesn't seem to be legitimate. The considered folk appraisals cannot be explained by KBDT. Indeed, such appraisals are not even compatible with the verdicts delivered by this theory. Far from providing a motivation for adopting a KBDT, such appraisals constitute a potential, quite bothersome problem for the theory.

It is not our goal to provide a diagnosis of why KBDT fails to explain the relevant data. However, our guess is that such problems originate from the specific decision-theoretic framework of the view. We can see this if we compare the respective explanations of rational action under conditions of uncertainty and risk provided by KAP and KBDT. Principles such as KAP allow deliberating from known propositions, where these include also propositions about what is probable. Now, there is a substantial difference between acting on (i.e., deliberating from) what we know, including what we know to be probable, and maximizing expected value given probabilities conditional on what we know. Some known propositions about probabilities may figure as appropriate premises in our practical reasoning but fail to count as good reasons for certain actions. For example, if we know the unfavorable odds of winning a lottery, we may well be in a position to use this knowledge as a reason for action and a premise in our deliberation, but this known fact may not in itself count as a *good* or *sufficient* reason to conclude to sell a ticket for a few cents. Things are very different when we consider the role of knowledge in a decision-theoretic framework: in that context knowledge only provides probabilistic distributions of possible

states of the world, which in turn affect expected values of outcomes by working as mere weights of possible values. Known propositions do not also count as normative considerations for or against certain actions. We suspect that this difference is responsible for the explanatory success of KAP and the failure of KBDT.[46]

As suggested by our previous comments, our point shouldn't be confused with the claim that it is always irrational to act on knowledge of probabilities. The latter claim is obviously false. It's undeniable that known probabilities of certain events can be good reasons for certain actions and attitudes (e.g., suspect that Jake stole your phone, take an umbrella if we learn that it's likely going to rain, and so on), though they are not good reasons for others (e.g., blame and accuse Jake of stealing your phone). These considerations could suggest a possible response to our objection to KBDT. One may argue that while KBDT provides wrong predictions about a range of cases, the theory is compatible with others in which it seems legitimate to act on what is probable given one's knowledge. An upholder of KBDT could then apply a 'divide and conquer' strategy and argue for a norm-pluralism where different norms are binding in different contexts.[47] While KAP can explain our assessments in problematic cases, KBDT could apply in all other situations in which it seems rational to act on statistical grounds.

We would like to point out two problems with this strategy. First, as argued by Hawthorne and Stanley (2008, pp. 581–584), KAP has the resources to explain cases of rational action under uncertainty without appealing to a KBDT. One possible maneuver is, for example, to appeal to knowledge of chances.[48] One may follow KAP's recommendations and input as a premise in one's reasoning the known fact that it is very likely that it will rain, which in itself can count as a reason to take an umbrella. In this perspective, implementing KBDT in our decision frameworks seems unnecessary and redundant,[49] since all the explanatory work could be done equally well or better by a single norm such as KAP. The appeal to a unique norm valid for all types of decisions would provide a more comprehensive and systematic, simpler, more straightforward and elegant explanation.

Second, even assuming that such a 'divide and conquer' strategy could be put to work and save the applicability of KBDT to some local areas of decision, we shouldn't lose sight of the broader dialectic of the debate. The whole point of the discussion in the present section was to challenge the presumption that KBDT can provide the best explanation of a range of knowledge-based ordinary folk appraisals commonly used in support of knowledge-action principles. In the literature we find that the main arguments for such principles appeal to cases such as lotteries, negligence and blame attributions. In contrast, cases of rational action in situations

[46] See Hawthorne and Stanley (2008, pp. 580–581) for further reasons to prefer principles such as KAP to a KBDT.

[47] Both Stanley (2007) and Moss (2018a, §9.2 and §9.3) express sympathy for a kind of norm-pluralism, even though their view is that different standards are operative across all contexts and figure in deliberation as *pro tanto* considerations.

[48] See also Stanley (2005, p. 10).

[49] And, we may add, impractical. See Gao (Unpublished manuscript) for a discussion of the difficulties of implementing decision theory in our ordinary decision-making.

of uncertainty have been generally considered problematic for knowledge-centered theories of rationality. Several philosophers have argued that actions can be rational even when based on mere credence or partial beliefs, also in the absence of knowledge of the relevant probabilities.[50] In our ordinary explanations of actions in situations of uncertainty we rarely appeal to knowledge, more often to other attitudes such as confidence. For example, we say that Mary took an umbrella because she was pretty confident that it was going to rain. In general, knowledge doesn't seem to be a necessary ingredient in these explanations. While upholders of knowledge-action principles have replies to this challenge,[51] they also acknowledge that explanations of rational action under uncertainty represent a potential problem for their views to the extent that their theories do not provide the only, or even the best, explanation of the relevant data. If this is the dialectic of the debate, we must acknowledge that a limited applicability of KBDT to action under uncertainty may well be consistent with some of our ordinary appraisals, but cannot provide any serious motivation for endorsing the theory. If the only data that KBDT can explain are those that can be explained equally well or better by other theories, the best-explanation strategy in support of KBDT fails.

4 Conclusion

In defending knowledge norms of action and decision from potential objections, Moss (2018a, p. 193) considers the interesting hypothesis that we should admit a plurality of norms governing rational action, each expressing a different ordinary sense in which we should act. Each norm would spell out consistent considerations figuring in rational deliberation. Amongst these norms, Moss includes maximizing expected value according to a knowledge-based decision theory.

The hypothesis of norm pluralism is fascinating and worth further investigation. However, we seriously doubt that this hypothesis alone can succeed in deflecting objections to KBDT and ensuring the legitimacy of the theory. This is because the real problem for KBDT doesn't concern its internal coherence, or its coexistence and potential conflicts with other norms governing human behavior. While we agree that one is not forced to choose amongst different normative theories, we should also avoid an unmotivated proliferation of norms. In particular, we should include a norm in our list only if it specifies some natural or important sense in which it would be better, rational or obligatory to act. As we argued in this paper, this doesn't seem to be the case for KBDT.

---

[50] See in particular Cresto (2010) and Schiffer (2007). For example, Cresto points out that "the reason I am carrying an umbrella is a primary probabilistic commitment, and not a belief, or a piece of knowledge, about a particular probabilistic claim. […] I shall focus exclusively on the more restricted contention according to which agents should not invoke probability claims as reasons unless they know that such claims are true. I believe there are grounds to think that this is a problematic demand. Thus, it seems perfectly right to treat a particular probabilistic commitment C as a reason for acting, without thereby requiring knowledge of C—or belief therein, for that matter" (2010, pp. 327–328).

[51] Hawthorne and Stanley (2008, §3, objection 2 and response).

The following table summarizes the conclusions we reached in this article:

	Intuitive rationality ascriptions	Guidance	3-options cases	Objective assessments (3rd pers., advice…)	Objective value maximization	K-based folk appraisals
Subjective DT	✗	✓	✓	✗	✗	✗
Objective DT	✗	✗	✗	✓	✓	✗
K-A principles	✗	✗	✗	✗	✗	✓
KBDT	✗(?)	✗	✗	✗	✗	✗

A way to summarize these results is that KBDT seems to lack all the advantages of alternative theories, while retaining their drawbacks. To use a metaphor, this theory is like an unlucky slalom skier who succeeds in missing all the gates. The theory is not subjective enough to guarantee guidance or avoid problems with three-options cases, but not sufficiently objective to account for ordinary standards of objective assessment (third-person, regret, advice) or express a useful notion of objective value maximization. Moreover, differently from other knowledge-action principles, the theory is even not able to account for ordinary folk appraisals of rational action based on knowledge. Given these premises, it is legitimate to ask why we should want to add to our list of norms one that fares worse than others on every considered score, collecting all the same types of problems without retaining any of the advantages.

Of course, our arguments are not meant to constitute a definitive refutation of KBDT. Rather, our aim in this paper was to put forward a challenge to upholders of the theory. While the contemporary debate surrounding KBDT has focused on technical issues, we think that more urgent questions for upholders of the view should be whether the theory is sufficiently well motivated, what its proper domain of application is, its specific utility and its advantages over alternative normative theories; in short, why we should want a knowledge-based decision theory.

References

Blome-Tillman, M. (2017). The knowledge norm of legal liability. In A. Carter, E. Gordon, & B. Jarvis (Eds.), *Knowledge first—approaches in epistemology and mind* (pp. 278–292). Oxford University Press.

Broome, J. (1995). *Weighing goods: equality, uncertainty and time*. Wiley.

Broome, J. (2013). *Rationality through reasoning*. Wiley.

Buchak, L. (2013). *Risk and rationality*. Oxford University Press.

Comesana, J. (2020). *Being rational and being right*. Oxford University Press.

Cresto, E. (2010). On reasons and epistemic rationality. *Journal of Philosophy, 107*(6), 326–330.

Dietz, C. (2018). Reasons and factive emotions. *Philosophical Studies, 175*(7), 1681–1691.

Dutant, J. (forthcoming). Knowledge-first evidentialism about rationality. In F. Dorsch & J. Dutant (Ed.), *The new evil demon problem*. Oxford University Press.

Fantl, J., & McGrath, M. (2002). Evidence, pragmatics, and justification. *Philosophical Review, 111*(1), 67–94.

Fassio, D. (2017). Is there an epistemic norm of practical reasoning? *Philosophical Studies, 174*(9), 2137–2166.

Fassio, D. (2021). Perspectivism, accessibility and the failure of conjunction agglomeration. *Ethics, 131*(2), 183–206.

Gao, J. (Unpublished manuscript). Should credence be sensitive to practical factors? A cost-benefit analysis.

Gao, J. (2017). Rational action without knowledge (and vice versa). *Synthese, 194*(6), 1901–1917.

Gardiner, G. (2018). Legal burdens of proof and statistical evidence. In D. Coady & J. Chase (Eds.), *The routledge handbook of applied epistemology* (pp. 179–195). Routledge.

Gerken, M. (2011). Warrant and action. *Synthese, 178*(3), 529–547.

Gibbons, J. (2006). Access externalism. *Mind, 115*(457), 19–39.

Gibbons, J. (2013). *The norm of belief*. Oxford University Press.

Goldberg, S. (2018). *To the best of our knowledge: Social expectations and epistemic normativity*. Oxford University Press.

Graham, P. A. (2010). In defense of objectivism about moral obligation. *Ethics, 121*(1), 88–115.

Greco, D. (2013). Probability and prodigality. In T. Gendler (Ed.), *Oxford studies in epistemology* (Vol. 4). Oxford University Press.

Harman, G. (1986). *Change in view. Principles of reasoning*. MIT Press.

Hawthorne, J. (2004). *Knowledge and lotteries*. Oxford University Press.

Hawthorne, J., & Srinivasan, A. (2013). Disagreement Without transparency: Some bleak thoughts. In D. P. Christensen & J. Lackey (Eds.), *The epistemology of disagreement: New essays* (pp. 9–30). Oxford University Press.

Hawthorne, J., & Stanley, J. (2008). Knowledge and action. *Journal of Philosophy, 105*(10), 571–590.

Heylen, J. (2016). Being in a Position to Know and Closure. *Thought: A Journal of Philosophy, 5*(1), 63–67.

Hoffman, D. D. (2019). *The case against reality: Why evolution hid the truth from our eyes*. Norton & Company.

Howson, C., & Urbach, P. (1993). *Scientific reasoning: The Bayesian approach* (2nd ed.). Open Court.

Hyman, J. (1999). How knowledge works. *Philosophical Quarterly, 50*(197), 433–451.

Jackson, F. (1991). Decision-theoretic consequentialism and the nearest and dearest objection. *Ethics, 101*(3), 461–482.

Jeffrey, R. (1965). *The logics of decision*. University of Chicago Press.

Jessica, B. (2008). Subject-sensitive invariantism and the knowledge norm for practical reasoning. *Noûs, 42*(2), 167–189.

Kaplan, M. (2009). Williamson's casual approach to probabilism. In P. Greenough (Ed.), *Williamson on knowledge* (pp. 122–139). Oxford University Press.

Kiesewetter, B. (2011). 'Ought' and the perspective of the agent. *Journal of Ethics and Social Philosophy, 5*(3), 1–24.

Kiesewetter, B. (2017). *The normativity of rationality*. Oxford University Press.

Kiesewetter, B. (2018). How reasons are sensitive to available evidence. In C. McHugh, J. Way, & D. Whiting (Eds.), *Normativity: Epistemic and practical* (pp. 90–114). Oxford Universty Press.

Kolodny, N., & MacFarlane, J. (2010). Ifs and oughts. *Journal of Philosophy, 107*(3), 115–143.

Levi, I. (1980). *The Enterprise of Knowledge*. MIT Press.

Lewis, D. (1981). Causal decision theory. *Australasian Journal of Philosophy, 59*, 5–30.

Littlejohn, C. (2009). Living with uncertainty. *Philosophical Books, 50*(4), 235–247.

Littlejohn, C. (2019). Being more realistic about reasons: on rationality and reasons perspectivism. *Philosophy and Phenomenological Research, 99*(3), 605–627.

Littlejohn, C. (2020). Truth, knowledge, and the standard of proof in criminal Law. *Synthese, 197*, 5253–5286.

Lord, E. (2018). *The importance of being rational*. Oxford University Press.

Mark, T. (2012). *Moral theory: An introduction*. Rowman and Littlefield Publishers.

Mason, E. (2013). Objectivism and prospectivism about rightness. *Journal of Ethics and Social Philosophy, 7*(2), 1–22.

Moss, S. (2018a). *Probablistic knowledge*. Oxford University Press.

Moss, S. (2018b). VIII—Moral encroachment. *Proceedings of the Aristotelian Society, 118*(2), 177–205.

Neta, R. (2009). Treating something as a reason for action. *Noûs, 43*(4), 684–699.

Papineau, D. (2019). The disvalue of knowledge. *Synthese*. https://link.springer.com/article/10.1007/s11229-019-02405-4#citeas.

Parfit, D. (2011). *On what matters* (Vol. 1). Oxford University Press.

Peterson, M. (2009). *An introduction to decision theory* (1st Edn). Cambridge University Press.

Regan, D. (1980). *Utilitarianism and co-operation*. Oxford University Press.

Rosenkranz, S. (2016). Being in a position to know and closure: Reply to Heylen. *Thought: A Journal of Philosophy, 5,* 68–72.

Rosenkranz, S. (2018). The structure of justification. *Mind, 127*(506), 309–338.

Ross, J. (2012). Rationality, normativity, and commitment. In R. Schafer-Landau (Ed.), *Oxford studies in metaethics* (Vol. 7). Oxford University Press.

Schiffer, S. (2007). Interest relative invariantism. *Philosophical and Phenomenological Research, 75*(1), 188–195.

Schulz, M. (2017). Decisions and higher-order knowledge. *Noûs, 51*(3), 463–483.

Skorupski, J. (2010). *The domain of reasons*. Oxford University Press.

Smith, H. M. (2010). Subjective rightness. *Social Philosophy and Policy, 27*(2), 64–110.

Smith, H. M. (2011). The prospective view of obligation. *Journal of Ethics and Social Philosophy., 5,* 1–9.

Srinivasan, A. (2015). Normativity without cartesian privilege. *Philosophical Issues, 25,* 273–299.

Stanley, J. (2005). *Knowledge and practical interests*. Oxford University Press.

Stanley, J. (2007). Replies to Gilbert Harman, Ram Neta, and Stephen Schiffer. *Philosophy and Phenomenological Research, 75*(1), 196–210.

Stephens, C. L. (2001). When is it selectively advantageous to have true beliefs? *Philosophical Studies, 105*(2), 161–189.

Stich, S. (1990). *The fragmentation of reason*. MIT.

Strawson, P. (1962). Freedom and resentment. *Proceedings of the British Academy, 48,* 1–25.

Sullivan-Bissett, E. (2017). Biological function and epistemic normativity. *Philosophical Explorations, 20,* 94–110.

Thomson, J. J. (1986). Imposing risks. In W. Parent (Ed.), *Rights, restitution, and risk* (pp. 173–191). Harvard University Press.

Unger, P. K. (1975). *Ignorance: A case for scepticism*. Oxford University Press.

Weatherson, B. (2012). Knowledge, bets, and interests. In J. Brown & M. Gerken (Eds.), *Knowledge ascriptions* (pp. 75–103). Oxford University Press.

Weisberg, J. (2011). Varieties of Bayeisanism. *Handbook of the History of Logic., 10,* 477–551.

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

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

Williamson, T. (2008). Why epistemology can't be operationalised. In Q. Smith (Ed.), *Epistemology: New philosophical essays*. Oxford University Press.

Zimmerman, M. J. (2006). Is moral obligation objective or subjective? *Utilitas, 18*(4), 329–361.

Zimmerman, M. J. (2008). *Living with uncertainty: The moral significance of ignorance*. Cambridge University Press.

Zimmerman, M. J. (2014). *Ignorance and moral obligation*. Oxford University Press.