

## Against Weatherson on How to Frame a Decision Problem

Thomas A. Blackson  
Arizona State University  
Philosophy Faculty  
School of Historical, Philosophical, and Religious Studies  
Coor Hall, 4th Floor | P.O. Box 874302, Tempe, AZ 85287-3902  
[blackson@asu.edu](mailto:blackson@asu.edu)

### Abstract

In “Knowledge, Bets, and Interests,” Brian Weatherson makes a suggestion for how to frame a decision problem. He argues that “the states we can ‘leave off’ a decision table are the states that the agent knows not to obtain.” I present and defend an example that shows that Weatherson’s principle is false. Weatherson is correct to think that some intuitively rational decisions wouldn’t be rational if states the agent knows not to obtain were not omitted from the outcomes in the decision problem. This, however, is not true of every rational decision. Weatherson’s principle for how to frame a decision problem is open to counterexample.

## Against Weatherson on How to Frame a Decision Problem

In “Knowledge, Bets, and Interests,” Brian Weatherson makes a suggestion for how to frame a decision problem. An action is rational, according to the standard view, just in case its expected value is at least as great as the expected values of all the alternatives. For each action open to the agent in the circumstances, there is a set of outcomes. Each outcome has a utility. In addition, each outcome has a probability that is conditional on the performance of the action. Given this framework, the expected value of an action is the sum of the utilities of the outcomes discounted by the conditional probabilities. This standard view, however, says nothing about how to determine the particular actions, outcomes, utilities, and probabilities that frame a decision problem in a given circumstance. This is a problem, and Weatherson offers a partial solution. He argues that “the states we can ‘leave off’ a decision table are the states that the agent knows not to obtain” (Weatherson 2012: 82). Thus, according to the principle Weatherson asserts, the states that the agent knows not to obtain may be excluded from the sets of outcomes.

Initially, Weatherson’s principle can seem plausible. After all, it is natural to think that deliberation is essentially a matter of bringing what one knows to bear on the question of what one should do.<sup>1</sup> Furthermore, some intuitively rational decision wouldn’t be rational if states the agent knows not to obtain were not omitted from the outcomes of the decision problem. Consider the following example. Suppose that I have a minor medical condition. It would be nice to be rid of it, but living with it is not too difficult. Suppose that there is a medication I can take to alleviate the condition. The medication carries a small risk of death, but the risk is small enough that I know the medication won’t kill me. It seems clear in this example that it would be rational for me to take the medication, and Weatherson’s principle helps explain why. Because I know the medication won’t kill me, the state in which the medication kills me is not an outcome

on the decision table. If it were, taking the medication wouldn't be rational. The probability of the medication killing me is low, but given the high negative utility of the medication killing me, not taking it is a better choice because living with the condition is not too difficult.

Weatherson's principle, however, makes the wrong prediction in a more general form of this example. Instead of there being one medication, suppose that there are two medications that I can take to alleviate the medical condition. Suppose that from the points of view of cost, convenience, etc., the medications are equivalent. Finally, suppose that although the risk of death is small for both medications, this risk is considerably smaller for the second medication. Since the risks are small enough, I know that neither medication will kill me. So, according to Weatherson's principle, neither of the states in which the medication kills me is an outcome on the decision table. Given this way of framing the decision problem, I should take one of the medications but which I take doesn't matter. This, however, seems wrong. Intuitively, it seems very clear that I should take the medication with the smaller probability of killing me.

One might try to reject this counterexample by arguing that I don't know that neither medication will kill me because I am aware that each carries a small but nonzero risk of death. But this line of reasoning is not promising. First of all, Weatherson himself says that "[i]f we can only leave off things that have probability 1, then decision theory would be useless; but it isn't" (Weatherson 2012: 82). Further, unless one is going to give in to skepticism (which would trivialize Weatherson's principle), there is good reason to say that I know that neither medication will kill me. In the one-medication example, it is natural to think that I know that the medication won't kill me and that my decision to take the medication wouldn't be rational if I couldn't omit that outcome from the decision problem. It is hard to see how the possibility of a second medication, whose probability of killing me is even smaller, could undermine this knowledge.

Moreover, given that the probability of the second medication killing me is considerably smaller than the probability of the first medication killing me, it is hard to see why I wouldn't know that the second medication won't kill me if I know that the first one won't kill me.

This argument may be stated more formally. Weatherson commits himself to the view that an agent can know that a state will not obtain even if there is a nonzero but small enough probability that the state will obtain. Let  $m$  be the probability that taking the first medication won't kill me, and let this probability be less than but close enough to 1 for Weatherson to preserve his commitment to the knowledge. Let the probability that the second medication won't kill me be less than 1 but much greater than  $m$ . Given this way of stipulating the two probabilities, Weatherson must admit that I know that the second medication won't kill me if I know that the first one won't kill me. In the one-medication example, it seems clear that I know that the first medication won't kill me. Moreover, since it is extremely implausible to think that the possibility of a second medication could undermine this knowledge, it follows that in the two-medications example I know that neither of the medications will kill me.

It is of no avail to respond to this argument by saying that I don't know that the medications won't kill me because the stakes are too high. First of all, even if knowledge is interest-relative, as Weatherson believes, the medication examples represent natural and clear situations in which the practical importance of the matter doesn't encroach on the knowledge of what will and will not happen. People typically realize that medications have side-effects, including death, but if the relevant probabilities are low enough, they conclude that these side-effects will not occur. To draw this conclusion, they use Statistical Syllogism:

**PROB** ( $x$  is  $B$  |  $x$  is  $A$ ) is high (low).

This is an  $A$ .

-----

This is (not) a  $B$ .<sup>2</sup>

On the basis of what the doctor has said, the agent knows that the probability of the medication killing a patient is low on the condition that the patient takes the medication as prescribed. It is low enough, given the practical importance of the matter to the agent, that the agent reasonably concludes that the medication will not kill him if he takes it as prescribed.

Secondly, Weatherson can't deny the knowledge in the two-medications example without undermining his own principle. In the one-medication example, it is natural to think that taking the medication is rational and that it would not be rational if I did not know that the medication won't kill me. My knowledge that the medication won't kill me is part of the background information that I rely on to frame the decision problem. In this way, the two-medications example is significantly different from the potential counterexamples that have appeared in the philosophical literature, such as Brown's affair case (Brown 2008), Reed's jellybean case (Reed 2010), Brown's surgeon case (Brown 2012), and Reed's rain case (Reed 2013). These examples consist in a direct appeal to a situation in which one is supposed to have the intuition that an agent must include a particular outcome in the decision table even though the agent knows that it will not obtain.<sup>3</sup> In contrast to these examples, the two-medications example is built on an instance of Weatherson's own principle. The two-medications example is a straightforward generalization of the one-medication example, and the one-medication example is a situation that shows that an instance of Weatherson's principle is true. If, in response to the two-medications

example, Weatherson were to insist that I don't know that the medications won't kill me, he would undermine the motivation for his principle that the one-medication example clearly provides. Instead, he should admit that his principle is open to counterexample.

The intuition that informs Weatherson's principle helps explain the rationality of many ordinary actions, but this intuition about how to frame a decision problem is not correctly formulated as the universal claim that "the states we can 'leave off' a decision table are the states that the agent knows not to obtain." The two-medications example shows that a more nuanced formulation is necessary. What this formulation is, however, is unclear given the current state of understanding of how to frame a decision problem in a particular circumstance. Decision problems must be framed in terms of some of the agent's cognitive states, but the principles that relate the elements of the decision problem to the agent's cognitive states have thus far proven difficult to formulate in a way that is precise and immune to counterexample.<sup>4</sup>

---

<sup>1</sup> For questions about whether knowledge is the right epistemic state in terms of which decision problems should be framed, see Brown 2012: 56-57.

<sup>2</sup> For a discussion of Statistical Syllogism, see Pollock & Cruz 1999: 229-234.

<sup>3</sup> Reed's jellybean case is an example. In this case, an agent is participating in a psychological experiment in which he receives nothing for not attempting to answer a test question, a jellybean for giving a correct answer to a test question, and an extremely painful electric shock for giving an incorrect answer to a test question. The agent is stipulated to be confident but not absolutely certain of the answer to the question of the year in which Julius Caesar was born. About this case, Reed expects his readers to share his intuition that although it is not rational to attempt an answer, the agent knows when Julius Caesar was born. "In this case it seems clear that, when you

---

weigh the meager reward against the prospect (however unlikely) of excruciating pain, it would not be rational to attempt an answer to the question. Does this show that you don't *know* when Caesar was born? Not at all. Your shortcoming is merely a lack of certainty when only something in that neighborhood would do. There is nothing odd about continuing to think of yourself as possessing knowledge in this situation..." (Reed 2010: 229; emphasis in original).

<sup>4</sup> I am grateful to Sabra Nuel and an anonymous reader for helpful comments on previous versions of this paper.

## Bibliography

Brown, Jessica. 2008. "Subject-Sensitive Invariantism and the Knowledge Norm for Practical Reasoning." *Nous* 42 (2): 167-189.

Brown, Jessica. 2012. "Practical Reasoning, Decision Theory and Anti-Intellectualism." *Episteme*, 9 (1): 43-62.

Pollock, J. & Cruz, J. 1999. *Contemporary Theories of Knowledge*, 2nd edition. Rowman & Littlefield Publishers.

Reed, Baron. 2010. "Defense of Stable Invariantism." *Nous* 44 (2): 224-244.

Reed, Baron. 2013. "Fallibilism, epistemic possibility, and epistemic agency." *Philosophical Issues* 23 (1): 40-69.

Weatherson, Brian. 2012. "Knowledge, Bets, and Interests." In *Knowledge Ascriptions*, ed. Jessica Brown and Mikkell Gerken. Oxford University Press, 75-103.