Knowledge Entails Dispositional Belief

David Rose (Rutgers) & Jonathan Schaffer (Rutgers)
Draft of April 2nd, 2012

Orthodoxy has it that knowledge entails belief, but the case of the unconfident examinee has been said (Woozley 1952: 155; Radford 1966) to show otherwise:

Unconfident examinee: Kate is taking a history test. She had studied carefully and has been doing well on all the questions so far. She has now reached the final question, which reads "What year did Queen Elizabeth die?" As Kate reads this question she feels relief, since she had expected this question and memorized the answer. But before Kate can pause to recall the date, the teacher interrupts and announces that there is only one minute left. Now Kate panics. Her grip tightens around her pen. Her mind goes blank, and nothing comes to her. She feels that she can only guess. So, feeling shaken and dejected, she writes "1603"—which is of course exactly the right answer.

Kate is said to know since she answers correctly, but to lack belief since she feels no confidence in her answer. So runs the leading challenge to the orthodox view that knowledge entails belief.

Unconfident examinee received sustained attention in the late 1960s and early 1970s, but the discussion ended in stalemate. Discussants were largely agreed that the matter turned on "when, in what situations, English speakers say or would say that someone knows or does not know" (Radford 1966: 5), but they disagreed about what English speakers would say, and—perhaps thinking that facts about the usage of ordinary language should be transparent—they lacked any further means to proceed past such disagreements. Thus the dominant response came to be Armstrong's (1969: 35) dismissive response: "I do not think that it is one of those clear cases that can be used as a test of a philosophical analysis." And so orthodoxy has stood.

Experimental philosophy provides new tools for resolving stalemates over what people will say. *One asks* (using psychological tools to devise the questions, and statistical tools to assess the answers). In this way Myers-Schulz and Schwitzgebel (*forthcoming*) have advanced the discussion, converting *Unconfident examinee* and four similar cases into empirical surveys. They (*forthcoming*: §3) report the following results: "A majority of respondents ascribed knowledge in our five scenarios, while only a minority ascribed belief," estimating that "perhaps about half" of their respondents had intuitions conforming to the view that knowledge does *not* entail belief. And so it may seem that orthodoxy should fall.

We applaud Myers-Schulz and Schwitzgebel for advancing the discussion in this way, and would only question their conclusion. We offer new empirical data supporting a new philosophical reply to the case, and conclude—shocking as it may sound for work in experimental philosophy—that the orthodox armchair view withstands empirical scrutiny.

Our reply is that Kate does believe—in the relevant sense—that Queen Elizabeth died in 1603. The epistemologically relevant sense of belief is not the *occurrent* notion of a thought consciously endorsed but rather the *dispositional* notion of information available to mind. Kate—though she lacks occurrent belief—still has dispositional belief since she has the information stored in mind from her studies and indeed draws on it to "guess" rightly. Her dispositional belief is merely *masked* by her temporary panic. We provide empirical support for this reply by reviewing Myers-Schulz and Schwitzgebel's results, and then presenting new variants of their studies which specifically elicit the

dispositional conception of belief (while removing biases: one of their five studies involves negative moral valence). We find that people's intuitions revert to conformity with the view that knowledge entails (dispositional) belief. In short, we redid Myers-Schulz and Schwitzgebel's studies to clarify the relevant dispositional notion of belief *and reversed their results*.

1. The Unconfident Examinee has Dispositional Belief

The case of *Unconfident examinee* represents the leading challenge to the orthodox idea that knowledge entails belief. We begin by clarifying the orthodox view and the dispositional notion of belief it involves, in order to explain why we think that *Unconfident examinee* fails as a counterexample: the unconfident examinee has dispositional belief.

1.1 The entailment thesis

Knowledge is widely thought to entail belief. This idea—labeled 'the entailment thesis' by Lehrer (1968: 491)—traces back at least to Plato's *Theatetus*, where knowledge is identified with true belief plus a *logos*. This idea is then codified in Gettier's (1963: 121) presentation of the traditional view as analyzing knowledge as justified true belief, preserved in the post-Gettier view that knowledge is justified true belief plus some fourth component, and retained even in Williamson's (2000: 41-8) primitivist break from the post-Gettier view.

The entailment thesis is widely endorsed because it seems intuitively plausible. For instance, if Moore knows that he has hands, then it seems to follow that Moore believes that he has hands. After all, how could Moore know that he has hands if he doesn't even believe it? The entailment thesis is also widely endorsed because it seems theoretically apt, reflecting an aspect of the idea that knowledge involves *fit* from mind to world. If the subject does not believe a given truth, then it seems as if her mind has thereby failed to fit the world in this respect, and she thereby does not deserve the credit for achieving knowledge of this point.

The entailment thesis may be regimented in various ways, but the simplest and most natural regimentation is via the following schema (Armstrong 1969: 21):

First pass: if s knows that p, then s believes that p

It is worth noting that the guiding idea behind the entailment thesis is more neutral than *First pass* might suggest. The guiding idea might be described—in admittedly vague terms—as the idea that where there is knowledge, the subject will have the relevant pro-attitude toward something relevantly connected to the object of knowledge. *First pass* adds a binary conception of knowledge as a relation between a subject s and proposition p. *First pass* also adds the ideas that the relevant pro-attitude is belief, that belief is a binary relation between a subject s and proposition s, and that the relevantly connected object of this attitude is the same as the object of knowledge. But the additions of *First pass* are largely peripheral to our discussion, and so we proceed from *First pass* for simplicity and definiteness.

¹ One might think that knowledge involves additional arguments such as a contrast proposition (Schaffer 2005), or that knowledge is not a relation to a proposition or some other representation of reality but rather to a fact or some other portion of reality itself (Vendler 1972).

² One might think that the relevant pro-attitude is acceptance rather than belief (Cohen 1989), or that belief involves additional arguments such as a degree argument (a view that traces back at least to Locke's *Essays*, Book 4, chs. 15-19), or that the relevant object of belief is related to but still different from the object of knowledge (for instance, Vendler (1978: 86) says of "the venerable formula: *x* knows that *p* iff *x* believes that *p*, *p*, and (something or other)" that he only asks for it to be amended to "*x* knows that *p* iff *x* believes that '*p*', etc.").

It is also worth noting that the guiding idea behind the entailment thesis may be less neutral than *First pass* might suggest. For the guiding idea also seems to involve the claim that the entailment in question is *analytic*. The entailment is not narrowly logical insofar as it depends on the interpretation of the 'knowledge' and 'belief' predicates, but it is not merely metaphysical (like the entailment from *being water* to *being H₂O*) in a way that is invisible to conceptual analysis. But these matters—and the vexed notion of analyticity they involve—are also largely peripheral to our discussion.

What is central to our discussion is a clarification of the notion of belief involved. The relevant conception of belief is not *occurrent belief* in the sense of a thought consciously endorsed, but rather *dispositional belief* in the sense of information available to mind. We thus see fit to amend *First pass* to make this explicit, and so work with:

Final pass: if s knows that p, then s dispositionally believes that p

In short: knowledge entails dispositional belief.

1.2 Occurrent and dispositional belief

The distinction between the occurrent and dispositional conceptions of belief—marked explicitly in *Final pass*—traces back at least to Ryle's (1949: 135) account of belief as "a propensity," and subsequent commentators—such as Campbell (1967)—who sought to make room for both Ryle's dispositional conception and an equally legitimate "episodic" conception. The basic contrast is between a thought consciously endorsed ("occurrent belief"), and information available to mind for endorsement ("dispositional belief"). We take it that both are perfectly legitimate conceptions, and that the English word 'believe' is moreover *polysemous* between an occurrent and a dispositional sense. 4

Occurrent belief is something like explicit judgment, involving the conscious endorsement of the content. It is the mental counterpart of overt assertion. Arguably it involves a distinctive phenomenology: there is something characteristic that it is like to enjoy an occurrent belief, and perhaps there is even something characteristic for each content occurrently believed.⁵ A normal human being has a relatively sparse handful of occurrent beliefs at any given time, and may have no occurrent beliefs at many times (such as when her mind goes blank, or when she is asleep).

But occurrent belief is clearly not the sense of belief relevant to the entailment thesis. If occurrent belief were required for knowledge one would know far too little. A normal human adult knows all sorts

_

³ Campbell (1967: 206) also argues that the occurrent ("episodic") conception is more fundamental than the dispositional conception, since the relevant dispositions include: "the tendency to react to some of the relevant situations with *episodic* belief." Price (1969) distinguishes the traditional "occurrence analysis" of belief in terms of an introspectible mental act, from the "modern" "dispositional analysis" in terms of overt behavior. Lycan 1988 (ch. 3) and Audi 1994 offer related pluralistic taxonomies. See Schwitzgebel 2010 (§§2.1-2.2) for a useful overview of these matters.

⁴ Though we think that 'believe' is polysemous, our discussion is compatible with thinking of 'believe' as fully ambiguous, or as having a single neutral meaning which is fleshed out in different ways in different contexts (Ludlow 2008), or even as only naturally having the occurrent meaning so long as a stipulative dispositional meaning can readily be created on the fly. Indeed we aren't even committed to thinking that 'believe' has an occurrent meaning. Our minimal commitment is just that (i) there are multiple meanings for 'believe,' (ii) Myers-Schulz and Schwitzgebel's studies did not uniformly elicit the epistemically relevant reading, and (iii) the techniques we used to re-do their studies did more to uniformly elicit the epistemically relevant reading. See §3.1 for further discussion.

⁵ In this vein, Horgan and Tienson (2002: 526) argue that not only is the occurrent belief state associated with a characteristic phenomenology, and not only are occurrent belief contents also associated with a characteristic phenomenology, but moreover this phenomenology partly *constitutes* both the state and its content.

of mundane propositions about her own personal life, current events, basic arithmetic, and various other topics at any given time (even when her mind goes blank; even when she is asleep). For instance, a normal human adult will have known that 7+5=12 since early childhood without interruption, while only occurrently believing this on a handful of scattered occasions. Indeed a normal human adult will know all sorts of basic arithmetic truths without having explicitly considered them even once.

Accordingly, the sense of belief relevant to the entailment thesis is not occurrent belief but dispositional belief. The abundance of mundane propositions that most human adults know at any given time—including the many truths of basic arithmetic—are believed in the sense of being available to mind for endorsement. This is a dispositional notion, not an occurrent notion. It involves no characteristic phenomenology, but merely a disposition to such phenomenology. A normal human being may well have infinitely many dispositional beliefs at any given moment.

Klein's (1999) defense of *infinitism* provides a vivid illustration of the importance of distinguishing occurrent from dispositional beliefs. Infinitism is the view that justification for a given proposition p requires a limitless sequence of reasons r1, r2, ... such that r1 is a reason for p, r2 is a reason for r1, ... The classic objection to infinitism—tracing back at least to Aristotle's *Posterior Analytics*—is that our finite minds could not possibly hold the infinitude of beliefs that would be required to possess propositional justification so understood. Klein (1999: 300) rebuts this classic objection by noting that it trades on the (irrelevant) occurrent conception of belief: "Humans have many beliefs that are not occurrent. It is the non-occurrent sense of 'belief' that the members of an infinite series of reasons might be subjectively available to S." Our finite minds cannot hold an infinite series of occurrent beliefs, but they can ground an infinite series of dispositional beliefs.

The example of infinitism is useful not only for showing the philosophical relevance of the distinction between occurrent and dispositional belief, but also for showing how easy it is to confuse these closely related notions. The classic objection to infinitism *seems* compelling. Not for nothing did it persuade Aristotle and virtually all other participants to the discussion prior to Klein. So it should not be too surprising if it turns out—as we suggest—that the participants to the discussion of the case of the unconfident examinee suffered the very same tempting confusion.

1.3 Why the unconfident examinee has dispositional belief

So far we have introduced the entailment thesis and clarified the relevant dispositional notion of belief. It is time to return to *Unconfident examinee*, which surfaces in skeletal form (targeting the claim that knowledge requires certainty) in Woozley (1952: 155): "[I]f being sure is necessary to knowing, then many candidates at viva voce examinations have been credited with knowledge which they haven't got." Radford (1966, 1970a, 1970b, 1972, 1988, 1990)—whose unconfident examinee "Jean" faces various questions about English history—is mainly responsible for fleshing out the case, extending it to challenge the claim that knowledge requires belief, and championing it against various replies.⁶

In order for the case to stand as a counterexample to *Final pass*, the following two claims must hold of the unconfident examinee:

Kate knows: Kate knows that Queen Elizabeth died in 1603 *Kate lacks dispositional belief*: Kate does not dispositionally believe that Queen Elizabeth died in 1603

We accept *Kate knows*, but deny *Kate lacks dispositional belief*, instead upholding:

⁶ Radford also managed to convince several others, including Black (1971: 154), Margolis (1973: 7), Mannison (1976: 139), Lewis (1996: 556), and Shope (2002: 53-55).

Kate dispositionally believes: Kate dispositionally believes that Queen Elizabeth died in 1603

So we claim that *Unconfident examinee* provides no counterexample to the entailment thesis once properly clarified as per *Final pass*.

What does Kate "believe"? To begin with, it is clear that Kate does not have the occurrent belief that Queen Elizabeth died in 1603. After all, Kate is unable to consciously endorse the thought. Her mind has gone blank. She feels as if she is guessing.

Indeed, notice that all the parts of the case suggestive of lack of belief (lack of felt confidence, a feeling of guessing) operate at the level of phenomenology, which only speaks to the issue of occurrent belief. Given that occurrent and dispositional belief are easily confused (§1.2), it is therefore unsurprising that people who are not explicitly guided to the relevant dispositional conception of belief will be guided by the prevalence of phenomenological description into operating with the irrelevant occurrent conception, and thereby intuit that Kate does not believe that Queen Elizabeth died in 1603. Such intuitions are perfectly reasonable. They just happen not to involve the relevant notion of belief, and so provide no support for *Kate lacks dispositional belief*.

The relevant question is whether Kate has the dispositional belief that Queen Elizabeth died in 1603. We uphold Kate has dispositional belief for two main reasons. First, Kate has the information stored in mind from her studies. It is lodged in her memory, and presumably the memory trace has not been completely and irrevocably destroyed by her momentary panic. Secondly, Kate does guess correctly, and presumably this is no accident. Indeed it seems as if her memory trace must still be not just present but actually operating in the background to guide her actions, even if she is unable in the moment to appreciate the fact. Putting these two reasons together—to the extent that it is useful to operate with the picture of a "belief box" in which various propositions are stored—we find it natural to think of Kate as having the proposition that Oueen Elizabeth died in 1603 lodged in her belief box throughout. She stored it there during her studies and is still unconsciously guided by it when she "guesses." Indeed we find it natural to imagine that—perhaps later that very day—Kate will recover from her panic and recall the information readily enough. She has the information stored in mind. She is merely temporarily blocked from accessing it normally.⁷

That said, there is a plausible—but ultimately misleading—rationale for upholding Kate lacks dispositional belief. The rationale is that dispositions are connected to simple counterfactuals which say that if a triggering condition for the disposition were to obtain, then the manifestation condition of the disposition would result. Indeed—until the early 1990s, after the discussion of the case of the unconfident examinee had largely died out—it was fairly standard to analyze dispositional statements in terms of such simple counterfactuals. The relevant simple counterfactual is *false* with Kate: indeed she is actually asked the question (a triggering condition for dispositional belief) but still fails to manifest occurrent

right answers and does so, when he does, not by fluke, but because he once learned them, or came into contact with them and they 'stuck'..." If the unconfident examinee turned out not to have the information stored in mind after all, she could not be said to remember and so could not be said to know (or at least she could not be said to know for anything like Radford's reasons).

⁷ In this yein, Radford (1990: 617) defends the claim that the unconfident examinee *knows* via the slogan: "He remembers, so he knows." He elaborates: "To say of Jean that he remembers is to say that he can give the relevant

⁸ The counterfactual analysis of dispositions traces at least to Goodman's (1955: 35-40) criticism of Carnap's extensional analysis, and to Ryle's (1949: 44) "inference ticket" conception of dispositional claims. Most historical participants to the discussion of the unconfident examinee were operating in a context in which dispositions were associated with simple counterfactuals, though we note that Woozley himself (1948) voiced early skepticism of the simple counterfactual account.

belief (the manifestation condition). So it might be thought that Kate does not possess dispositional belief after all. That is, it might be thought—on the basis of the simple counterfactual analysis of dispositions—that Kate does not really have the information available to mind, on grounds that she fails to actually access this information when asked.

Yet it is now widely accepted that the link between dispositions and counterfactuals is complex, in that there are various ways in which the triggering condition for the disposition may obtain without the manifestation condition resulting. One salient way in which the triggering condition may obtain without the manifestation condition resulting—which emerges in Johnston 1992 (p. 223)—is that extrinsic factors may *mask* the disposition. Johnston gives the example of a fragile glass which is struck (triggering condition) but does not shatter (no manifestation) due to careful packaging (the mask). In the case of the unconfident examinee, Kate is asked the question (triggering condition) but cannot properly recall her view on the matter (no manifestation) due to her panic (the mask). So we think that the case of the unconfident examinee is a case where the disposition is present (for the reasons given above) but merely masked. The falsity of the simple counterfactual does not establish *Kate lacks dispositional belief*.

Indeed we think—continuing to operate with the image of "the belief box"—that the intrinsic structure of Kate's belief box is apt for holding the belief. It is only her recall capacities—the pipeline from the belief box to consciousness—which are temporarily clogged by her flood of panic. Of course we do not need the belief box image to make this point. It is enough that Kate's belief system (however it is best understood) is intrinsically apt for dispositional belief. So further consideration of why the simple counterfactual fails turns out to actually provide a third reason in favor of *Kate has dispositional belief*: the case fits the pattern of a masked disposition.

Putting this together: not only have we given three reasons to support *Kate has dispositional belief*, we have also provided a two-fold explanation for why people might have wrongly thought that *Unconfident Examinee* was a counterexample to the entailment thesis, involving independently attested confusions. First, some people may have confused occurrent belief with dispositional belief. This is a confusion that afflicted the very best of philosophers for centuries in the case of infinitism. Secondly, some people may have rejected *Kate has dispositional belief* due to a reliance on an overly simplistic counterfactual analysis of dispositions. This is a confusion that ran rampant at the very time in which the case of the unconfident examinee was discussed.

So we conclude that *Unconfident examinee* fails as a counterexample to the entailment thesis. Of course there may be other counterexamples or other reasons to reject the entailment thesis. Though for the record we regard the case of the unconfident examinee as the main extant challenge to the entailment thesis extant—Stout (2006: 166) calls it "the supposed counterexample"—and so think that disarming this supposed counterexample constitutes the main defense currently required of the entailment thesis.⁹

_

⁹ The other main objection to the entailment thesis found in the literature begins from the observation that people say things like "I don't *believe* it, I *know* it," and concludes that knowledge and belief must be *incompatible*. This objection is sometimes traced to Plato in *Republic* (who speaks of knowledge and *doxa* as incompatible), and sometimes to Moore in *Commonplace Book* (who comments on a knowledge-compatible and incompatible sense of "believe"), and is at any rate defended in Prichard 1950. But this objection has virtually disappeared from the contemporary literature, for good reason. With the advent of Gricean pragmatics it came to be widely agreed that "I don't *believe* it, I *know* it" should not be understood as opposing belief and knowledge, any more than "I don't have *some* of the money, I have *all* of the money" should be understood as opposing some and all. See Harrison 1963, Lehrer 1968, and Armstrong 1969 for further discussion along these lines.

We note that our reply to *Unconfident examinee* is, as far as we are aware, a new reply. Discussion of the case largely died out by the mid 1970s. Defenders of the entailment thesis to date have generally taken the position—opposite to ours—that since Kate does not "believe" she thereby cannot really know (Jones 1966, Lehrer 1968, Sorensen 1982, Steup 2006, and Stout 2006). Or they have taken the (compatible) position that the case is too unclear to topple an independently plausible plank of orthodoxy (Armstrong 1969, Lehrer 1974, Dartnall 1986, and Williamson 2000). Indeed to our knowledge only Cohen (1966: 11) and Armstrong (1969: 35) have claimed that Kate both knows and believes, but neither uses the occurrent/dispositional distinction to back the claim that Kate believes.

Though our reply is new, we share the attitude of most participants to the discussion in treating our key claim—*Kate has dispositional belief*—as having empirical consequences concerning what ordinary speakers will say (Radford 1966: 5, Jones 1966: 19). In particular, we think that *if one can elicit the dispositional conception of belief* (without triggering any cognitive biases or other sources of performance error), then people will say that Kate believes that Queen Elizabeth died in 1603. In other words, we predict that our conceptual competence with dispositional belief will *show up* in the behavior of ordinary speakers, as long as no performance errors intervene. And we follow the lead of Myers-Schulz and Schwitzgebel (and experimental philosophers generally) in putting our empirical claim to the empirical test.

2. Myers-Schulz and Schwitzgebel's Results

Myers-Schulz and Schwitzgebel (*forthcoming*) have advanced the discussion of the entailment thesis by converting *Unconfident examinee* and four similar cases into empirical surveys. We now turn to a discussion of their results, in order to set the stage for a discussion of our own results, which we see as largely founded upon—and in some respects building on, while in other respects building over—their results.

2.1 Five studies, five worries

Myers-Schulz and Schwitzgebel (*forthcoming*) develop five separate empirical surveys designed to test the entailment thesis against counterexamples like *Unconfident examinee*. (The interested reader may find all of their vignettes, as well as our variants, in the appendix.) Aggregating responses to these diverse cases, they (*forthcoming*: §3) report: "Across the five main scenarios, 77% of respondents attributed knowledge and 41% attributed belief. These percentages are statistically significantly different from each other and, in both cases, from 50%." Though of course aggregating results from diverse studies is methodologically questionable, especially when it is not obvious that the studies exhibit the same phenomena. Accordingly they immediately move to the level of the five individual studies. At this level they report a statistically significant difference between knowledge attribution and belief attribution in only three of their five studies (their unconfident examinee, prejudiced professor, and freaked-out moviewatcher studies). Since we are interested in defending the entailment thesis, we hereby ignore the two

_

¹⁰ Indeed Radford (1988: 496), revisiting the discussion, spoke of it as "this ancient issue" and (1988: 497) spoke of a temptation—which he did warn was "too quick"—to stop "with an acid remark that finding the debate increasingly sterile I dropped out of it years ago, and things have not changed."

¹¹ Cohen (1966: 11-12) speaks of the intuition that Kate does not believe as arising from a "short-term criteria for believing" which looks at Kate's behavior in the moment rather than a "long-term criteria" which looks at her behavior across a longer span. Armstrong's explanation of the intuition that Kate does not believe comes from his (1969: 34) suggestion that perhaps Kate has *both* the belief that Queen Elizabeth died in 1603 *and* the conflicting belief that Queen Elizabeth did not die in 1603, and that people confuse the truth that Kate's believes that Queen Elizabeth did not die in 1603 with the falsehood that Kate lacks the belief that Queen Elizabeth died in 1603. ¹² Myers-Schulz and Schwitzgebel characterize all five of their cases as exhibiting "in-between" belief. As will emerge below we think that their cases are quite diverse (for instance, some of their cases exhibit a temporary blocking of a single standing belief, while other cases exhibit an enduring conflict between two standing beliefs), and call for diverse explanations.

studies (their absent-minded driver and self-deceived husband studies) for which no statistically significant difference between knowledge attribution and belief attribution was found, as not even presenting a *prima facie* challenge.

So far we have done little to mitigate Myers-Schulz and Schwitzgebel's challenge to the entailment thesis. We have only noted (and we take it that they would agree) that they have not *five* but *three* empirically supported *prima facie* counterexamples to the entailment thesis. *One* counterexample would of course suffice to refute the thesis.

That said, we note five worries about their results, which will serve to set our agenda. The first worry is that Myers-Schulz and Schwitzgebel use a between-subjects design throughout. That is, in each of their studies, they had one group of participants who were asked the knowledge question, and a second distinct group of participants who were asked the belief question. While a between-subjects design is of course methodologically legitimate, it is still generally inferior to a within-subjects design in which each participants gets to answer *both* the knowledge and the belief questions. For only in a within-subjects design can one directly determine whether any one person would be willing to ascribe knowledge without belief. And only in a within-subjects design can ordering effects be considered (Usually one only resorts to a between-subjects design when one worries that seeing one question will bias behavior on the other.) We would like to see these results replicated using a within-subjects design.

Secondly, Myers-Schulz and Schwitzgebel do not consider perspective. As has emerged from discussion of claims to know what is false ("The medievals knew that the earth was flat"), there is a systematic ability to use language *projectively*, speaking "through the mouths" of third parties. ¹³ Accordingly it is not clear that the attributions made reflect a given participant's own view of who knows and who believes what, or if they reflect this participant's view of what a third party (e.g. Kate herself) would say. Indeed one might worry that in *Unconfident examinee*, the majority of participants hold from their own view that Kate both knows and believes, hold from the view of what Kate would say that Kate neither knows nor believes, and then read the belief probe but not the knowledge probe projectively. We would like to see these results replicated using probes that explicitly mark perspective.

Thirdly—and crucially for our discussion of their unconfident examinee and freaked-out movie-watcher cases—Myers-Schulz and Schwitzgebel do not consider whether their belief probes are being read occurrently or dispositionally. That is, their participants are simply asked whether a character in the vignette believes a given proposition. Nothing is done to ensure that the relevant dispositional conception of belief is elicited, rather than the irrelevant occurrent conception. Indeed, even if it was a matter of chance which conception they elicited in a given participant, this would still yield the prediction that knowledge attributions would tend to outpace belief attributions, insofar as participants for whom the occurrent conception of belief was elicited would tend to attribute knowledge without (occurrent) belief. We would like to see these results replicated using probes that specifically elicit the relevant dispositional conception of belief.

1

¹³ See Buckwalter (*manuscript*) for results on the factivity of knowledge that turn on resolving perspective. For related results concerning reference in Kripke cases, see Sytsma and Livengood (2011). This worry is thus a known problem worth considering.

 $^{^{14}}$ Suppose that it was a matter of 50% chance which conception they elicited in a given participant. And also suppose—just to work with simple albeit unrealistic numbers—that 100% of subjects intuit that knowledge entails dispositional belief, and that 100% of subjects intuit that Kate lacks occurrent belief. Then one predicts that if x% of participants say that Kate knows, then x/2% of participants would say that Kate believes. This is actually not a poor fit for the 87%-37% spread that Myers-Schulz and Schwitzgebel actually report! Of course we are not suggesting that this is exactly what happened (the hypothesis that 100% of subjects intuit something is almost always unrealistic), but only pointing out one way in which failing to consider the occurrent/dispositional distinction calls Myers-Schulz and Schwitzgebel's interpretation of the data into question.

Fourthly—and crucially for our discussion of their prejudiced professor case—Schwitzgebel and Myers-Schulz also allow complicating considerations of moral valence into one of their vignettes. That is, in their prejudiced professor case (*appendix*, V5) participants are asked to consider a character who is prejudiced, and whom they will doubtlessly wish to *condemn*. It has been long known—ever since Knobe's (2003) pioneering work—that such matters of moral valence have profound implications for what we will say about a given agent's mental states. While the jury is still out on the best explanation of this "Knobe effect," at least one leading approach views moral valence effects as *biasing*, triggering a cognitively distorting desire to lay blame on individuals who arouse disapprobation (Alicke 1992). While we do not wish to take a stand on the best explanation of the Knobe effect here, it seems to us that their prejudiced professor case as presented involves a potentially biasing factor. We would like to see these results replicated with a revised vignette with neutral moral valence.

Fifth and finally, Myers-Schulz and Schwitzgebel do not explain why only three of their five cases showed a statistically significant difference between knowledge attribution and belief attribution. What explains the *diversity* of their results? The entailment thesis they oppose seems to predict that knowledge ascription should never outpace belief ascription. But Myers-Schulz and Schwitzgebel's (*forthcoming*: §4) own *capacity-tendency account* seems equally to predict uniformity in the opposite direction. Their capacity-tendency account has it that knowledge involves a clear capacity to access the right information, but that belief requires a bit more, namely a clear tendency to consistently deploy the information in action. (This has the nice consequence that knowledge *typically* comes along with belief, even while allowing occasional cases of knowledge without belief.) All of their cases seem equally to involve subjects with a clear capacity to access the right information but without a clear tendency to consistently deploy the information in action. (Indeed this seems to have been their guiding idea in developing their five cases.) We would like this diversity of results explained.¹⁵

Putting this together, we are worried about the following points:

- Whether Myers-Schulz and Schwitzgebel's results can be replicated using a within-subjects design
- Whether their results can be replicated with explicitly marked perspective
- Whether their results for the unconfident examinee and freaked-out movie-watcher cases can be replicated specifically eliciting a dispositional reading of 'believe'
- Whether their results for their prejudiced professor case can be replicated with neutral moral valence
- Why a statistically significant difference between knowledge and belief ascription only showed up in three of their five studies

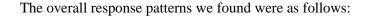
We hasten to add that, of these five worries, we regard the third and fourth together as the most worrisome. By our lights, the other worries still leave Myers-Schulz and Schwitzgebel with at least a *prima facie* case against the entailment thesis. (The first and second worries merely concern potential problems not yet shown to have arisen, and the fifth worry only concerns whether their own alternative capacity-tendency account fits the data any better.) Indeed we mention the first and second worries only because—as we will now show—these worries can be empirically *overcome*.

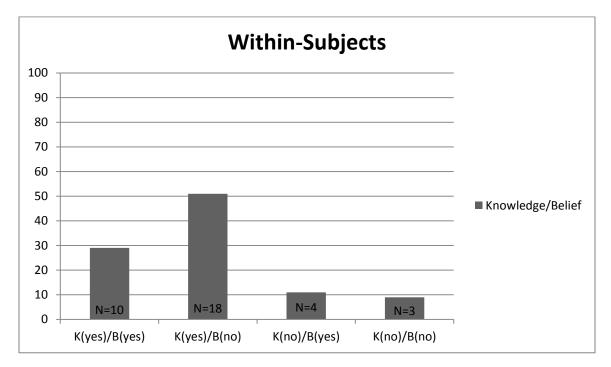
_

¹⁵ Myers-Schulz and Schwitzgebel *could* try to explain the diversity of their results from their capacity-tendency account, for instance by arguing that the clarity of the tendency to consistently deploy the information is varying between their cases. We are not saying that there is no possibility of using their capacity-tendency account to explain the diversity of their results. We are only saying that they do not in fact make any such attempt, and that it is not obvious that any such attempt would succeed.

2.2 Resolving the first worry: within-subjects replication

In order to address our first worry of whether Myers-Schulz and Schwitzgebel's results can be replicated using a within-subject design, we re-ran their version of *Unconfident examinee* using a within-subjects design. We used their original vignette (V1) and their knowledge and belief probes together, counterbalancing the order in which participants were given these probes, so that we could also see whether there was an ordering effect.





As can be seen, the primary response to the case of the unconfident examinee (just over 50%) is to *both* ascribe knowledge and deny belief. While Myers-Schulz and Schwitzgebel had estimated indirectly that "perhaps about half" of their respondents had intuitions conforming to the denial of the entailment thesis, we were able to substantiate this estimate directly.

Moreover, because we counterbalanced the order in which participants received the knowledge and belief probes, we were able to look for order effects. We found no evidence of order effects on either knowledge¹⁶ or belief¹⁷ attributions.

We were also able to look specifically at subgroups, including the subgroup that ascribed knowledge, and the subgroup that denied belief. We were thus able to ask, of the subgroup that ascribed knowledge (e.g. filtering out skeptics), what percentage of that subgroup also denied belief. We found that the majority (64%) of this subgroup also denied belief. And we were thus able to ask, of the

Among those who ascribed knowledge, the proportion of those who denied belief (18) is greater than the proportion of those who ascribed belief (10), p=.002, binomial test proportion=.36.

10

¹⁶ There were no significant order effects on attributions of knowledge, Fisher's Exact Test, p=.692.

¹⁷ There were no significant order effects on attributions of belief, Fisher's Exact Test, p=.732.

subgroup that denied belief, what percentage of that subgroup also ascribed knowledge. We found that the majority (86%) of this subgroup also ascribed knowledge. ¹⁹

Putting this together, we found—in ways that significantly *strengthen* Myers-Schulz and Schwitzgebel's original between-subjects results—the following four results from our within-subjects design:

- The primary response (just over 50%) was to both ascribe knowledge and deny belief
- No order effects were detected concerning the knowledge and belief probes
- Among the subgroup that ascribed knowledge, the majority (64%) denied belief
- Among the subgroup that denied belief, the majority (86%) ascribed knowledge

We thus regard our first worry as empirically resolved.

2.3 Resolving the second worry: explicit perspective replication

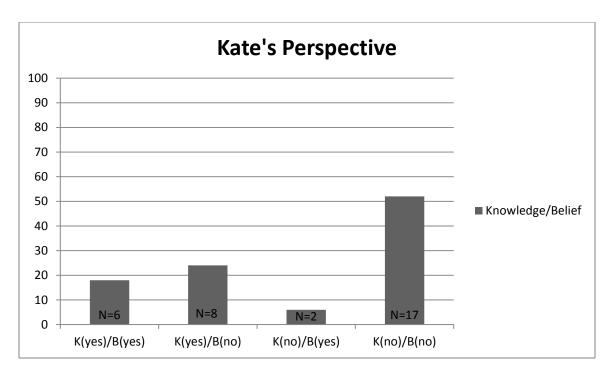
In order to address the second worry of whether Myers-Schulz and Schwitzgebel's results can be replicated with explicitly marked perspective, we also re-ran their version of *Unconfident examinee* (V1) with revised probes, specifically asking one group of participants the questions of "Did Kate actually know?" and "Did Kate actually believe?" (*participant perspective*), and asking a second group of participants the questions of "Did Kate think that she knew?" and "Did Kate think that she believed" (*Kate's perspective*). This also allowed us to consider the independently interesting question of how participants will view things from Kate's perspective.²⁰

We begin with Kate's perspective, both as an independently interesting matter, and in order to see if Myers-Schulz and Schwitzgebel's results might be arising from participants adopting this perspective. Our own intuitions were that, speaking from Kate's perspective, we would deny both knowledge and belief (after all, Kate feels as if she is only guessing). In accord with our intuitions, we found that the majority of participants (52%) denied both knowledge and belief from Kate's perspective. Indeed only 24% ascribed knowledge but denied belief:

_

¹⁹ Among those who denied belief, the proportion of those who ascribed knowledge (18) is greater than the proportion of those who denied knowledge (3), p<.001, binomial test proportion=.86.

²⁰ In using probes of this sort to help explicitly mark perspective, we follow the lead of Sytsma and Livengood (2011) and Buckwalter (*manuscript*). For instance, Buckwalter compares probes of the form "[*character*] really knows" and "[*character*] thought she knew" to help confirm that the folk view knowledge as factive. We take Buckwalter's plausible results with these probes as some confirmation that they are marking perspective properly.



We found no evidence of order effects on either knowledge²¹ or belief²² attributions from Kate's perspective.

Looking a bit more closely into our data on Kate's perspective, we found that, among the subgroup that ascribed knowledge, the proportion of those who denied belief (57%) is not statistically different from the proportion of those who ascribed belief (43%).²³ And we found that, among the subgroup that denied belief, the proportion of those who denied knowledge (68%) was greater than the proportion of those who ascribed knowledge (32%).²⁴

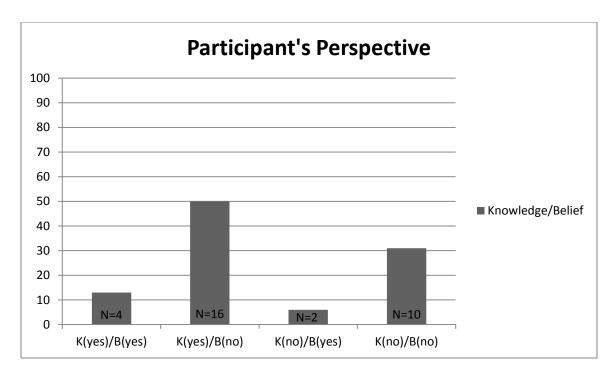
So far it seems as if Myers-Schulz and Schwitzgebel's results cannot plausibly be arising solely from participants taking Kate's perspective. Though this leaves open whether their results might plausibly be arising solely from participants taking their own perspective, or taking a mixture of perspectives, or even between shifting perspectives between the knowledge and belief probes (as we speculated in §2.1). So we turn to participants specifically taking their own "actual" perspective. We found that this subgroup behaved in ways quite similar to Myers-Schulz and Schwitzgebel's estimate based on their group of participants, and in particular found that the majority (50%) ascribed knowledge while denying belief:

²¹ There were no significant order effects on attributions of knowledge, Fisher's Exact Test, p=.080.

There were no significant order effects on attributions of belief, Fisher's Exact Test, p=.225.

²³ Among those who ascribed knowledge, there is no statistically significant difference between the proportion of those who denied belief (8) and the proportion of those who ascribed belief (6), p=.211, binomial test

proportion=.43. ²⁴ Among those who denied belief, the proportion of those who ascribed knowledge (8) was less than the proportion of those who denied knowledge (17), p<.001, binomial test proportion=.32.



We (yet again) found no evidence of order effects on either knowledge²⁵ or belief²⁶ attributions.

Looking a bit more closely into our data on the participant's perspective, we found that, among the subgroup that ascribed knowledge, 80% denied belief.²⁷ And among the subgroup that denied belief. 62% denied knowledge.²⁸

Our data thus suggest that Myers-Schulz and Schwitzgebel's results do not turn on any perspective shifting. Whether or not their original results should be interpreted as arising from the participant's perspective, it seems that their sort of result is sufficiently robust to be replicated even with explicitly marked perspective. Thus we found—in ways that significantly strengthen Myers-Schulz and Schwitzgebel's original results which do not mark perspective—that the primary response (against just around 50%) is to both ascribe knowledge and deny belief, even when perspective is marked. We thus regard our second worry as empirically resolved.

So far we have strengthened Myers-Schulz and Schwitzgebel's case against the entailment thesis. But of course we have three further worries (§2.1). And, when we come to our third worry—specifically eliciting a dispositional reading of 'believe'—we will find a very different story, as predicted in §1.

3. Confirming Entailment

Our story so far has had two main strands. In §1 we argued from the armchair that the unconfident examinee has dispositional belief, and thus constitutes no counterexample to the entailment thesis properly understood. Though we offered an empirical prediction: if one can elicit the dispositional conception of belief (without triggering performance errors), then people will say that Kate believes that

²⁵ There were no significant order effects on peoples attributions of knowledge, Fisher's Exact Test, p=.076.

²⁶ There were no significant order effects on peoples attributions of belief, Fisher's Exact Test, p=.1.000.

²⁷ Among those who ascribed knowledge, the proportion of those who denied belief (16) is greater than the proportion of those who ascribed belief (4), p<.001, binomial test proportion=.20.

Among those who denied belief, the proportion of those who ascribed knowledge (16) was greater than the

proportion of those who denied knowledge (10), p=.013, binomial test proportion=.38.

Queen Elizabeth died in 1603. Then in §2 we reviewed Myers-Schulz and Schwitzgebel's empirical results, and noted various worries, including that they failed to specifically elicit the relevant dispositional reading of 'believe.' We are now ready to tie these strands together. Our primary point is that when the Myers-Schulz and Schwitzgebel studies are re-done to elicit the relevant dispositional reading of 'believe,' people's intuitions revert to conformity with the view that knowledge entails belief, exactly as we predicted. In short, we redid Myers-Schulz and Schwitzgebel's studies to clarify the relevant dispositional notion of belief *and reversed their results*, thereby empirically *confirming* the thesis that knowledge entails dispositional belief.

3.1 Confirming the third worry: eliciting a dispositional reading of 'believe' In order to address the third worry of whether Myers-Schulz and Schwitzgebel's results can be replicated specifically eliciting a dispositional reading of 'believe,' we re-did their unconfident examinee and freaked-out movie-watcher cases in a variety of ways, each designed to specifically elicit a dispositional conception of belief. It is of course not obvious that any given design is really eliciting a dispositional conception of belief. Accordingly we chose to use three different sorts of designs, each of which might reasonably be thought to elicit a dispositional conception of belief. Our thought was that if our results are uniform across such designs, then that would buttress the claim that these designs are uniformly successful in eliciting a dispositional conception of belief.

In our first design, we re-did Myers-Schulz and Schwitzgebel's unconfident examinee (V1) and freaked-out movie-watcher (V3) vignettes, but modified the belief probe to include a parenthetical clarification of the sense at issue. For instance, instead of simply asking if Kate believes that Queen Elizabeth died in 1603, we asked:

Did Kate still believe (in the sense that she still held the information in her mind even if she could not access it) that Queen Elizabeth died in 1603?

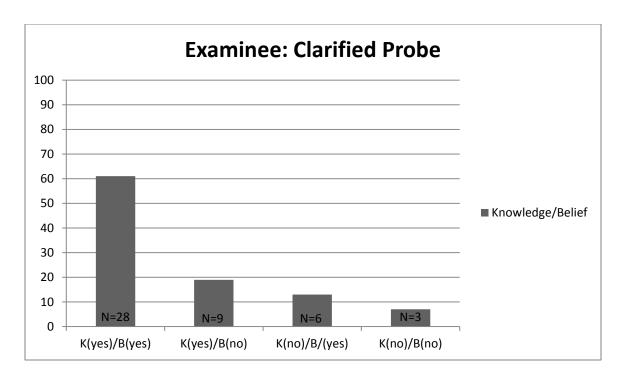
We thought of this as one reasonable attempt to specify the relevant dispositional sense of 'believe' (subject to the concern that some noise might be generated due to some participants being baffled).

So we ran their version of the unconfident examinee case (V1), retaining a within-subjects design (§2.2) and counterbalancing the order of the questions. Again we found no significant order effects on people's attributions of knowledge²⁹ or belief.³⁰ But the overall pattern of the responses changed drastically: the majority response (61%) became to attribute *both* knowledge and belief, just as we predicted in §1:

There were no significant order effects on people's attributions of belief df=2, X^2 =4.880, p=.087.

14

²⁹ There were no significant order effects on peoples attributions of knowledge, df=2, X^2 =.309, p=.857.



Among the subgroup that ascribed knowledge, the majority (76%) also ascribed belief.³¹ And among the subgroup that ascribed belief, the majority (82%) ascribed knowledge.³² And so it begins to seem as if specifically eliciting a dispositional reading of 'believe' might reverse Myers-Schulz and Schwitzgebel's results.

We then tried the same sort of manipulation with Myers-Schulz and Schwitzgebel's case of the freaked-out movie-watcher (V3), replacing their simple belief probe with the clarified probe:

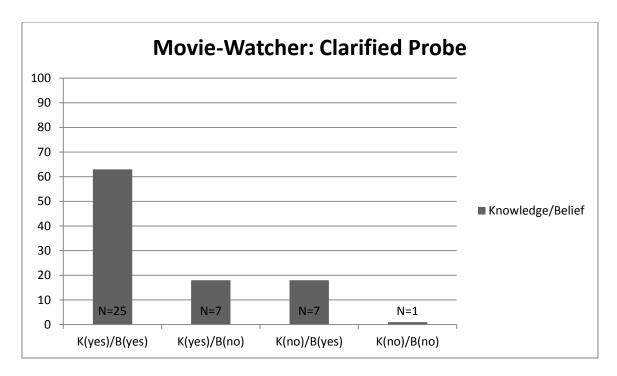
Did Jamie still believe (in the sense that she still held the information in her mind even if she could not access it) that only water would come out of the sink faucet?

We found a similar pattern of results, finding no order effects on people's attributions of knowledge³³ or belief.³⁴ and finding that the majority response (63%) was to attribute both knowledge and belief:

³¹ Among those who ascribed knowledge, the proportion of those who also ascribed belief (28) is greater than the proportion of those who denied belief (9), p<.001, binomial test proportion=.24.

32 Among those who ascribed belief, the proportion of those who also ascribed knowledge (28) is greater than the

proportion of those who denied knowledge (6), p<.001, binomial test proportion=.18. There were no significant order effects on people's attributions of knowledge, df=2, X^2 =1.521, p=.467.



Among the subgroup that ascribed knowledge, the majority (78%) also ascribed belief.³⁵ And among the subgroup that ascribed belief, the majority (78%) also ascribed knowledge.³⁶ And so it continues to seem as if specifically eliciting a dispositional reading of 'believe' might reverse Myers-Schulz and Schwitzgebel's results. Or at least, for two of the three studies on which Myers-Schulz and Schwitzgebel found evidence against the entailment thesis, it seems as if eliciting a dispositional reading of 'believe' does reverse their results.

Of course one might worry that our revised probe, with its parenthetical clarification of the sense of belief at issue, is not really being read as we intend.³⁷ Or relatedly—as Joshua Knobe (*personal communication*) brought to our attention—one might worry that our revised probe is being read in a purely stipulative way, not necessarily connected to any naturally occurring sense of 'believe.' We think that these are both perfectly legitimate worries. Our response is to see if we can replicate our previous results using a different design for eliciting a dispositional conception of belief, with the idea that uniform results would buttress the claim that these designs are uniformly successful in eliciting a dispositional conception of belief.

Accordingly we returned to the simple belief probes that Myers-Schulz and Schwitzgebel used, and instead complicated the accompanying vignettes to include discussion of a person said to be *asleep*. Our idea was that—since sleeping people do not typically have occurrent beliefs—we could use attributions of belief to a sleeping person to induce a context in which only the dispositional conception of belief could sensibly be taken to be in play. We note that this is a fairly subtle manipulation, and we had

-

³⁵ Among those who ascribed knowledge, the proportion of those who also ascribed belief (25) is greater than the proportion of those who denied belief (7), p<.001, binomial test proportion=.22.

³⁶ Among those who ascribed belief, the proportion of those who also ascribed knowledge (25) is greater than the proportion of those who denied knowledge (7), p<.001, binomial test proportion=.22.

³⁷ If 'believe' really is polysemous between an occurrent and a dispositional sense, we don't see how else our parenthetical clarification could be read. But (as mentioned in §1.2) we are open to other hypotheses, including ones on which underspecified lexical meanings might get specified "on the fly." That would allow that our parenthetical clarification might actually create a *third sense* of 'believe' that is neither the occurrent nor the dispositional sense.

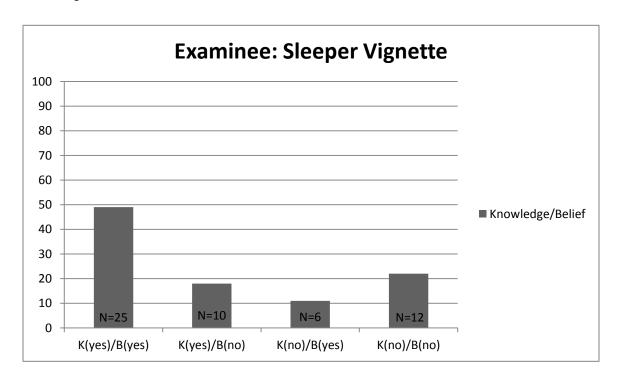
our doubts as to whether it would work. We are simply going to ask people—as Myers-Schulz and Schwitzgebel did—whether or not Kate believes that Queen Elizabeth died in 1603 (and likewise whether or not Jamie believes that only water would come out of the sink faucet)—and rely solely on the previous discussion of a sleeping person to reverse the results.

With that in mind we returned to *Unconfident examinee* and added a character—Dave—into the story who had studied just as much as Kate but who had slept through his alarm clock and was currently sleeping through the exam (V2). We then (while retaining a within-subjects design, and counterbalancing the order of the knowledge and belief probes) asked people about what Dave believed before asking about what Kate believed:

Did Dave (despite being asleep) believe that Queen Elizabeth died in 1603?

We thought of this as one reasonable indirect attempt to elicit a dispositional conception of belief when thinking about Kate, insofar as the dispositional conception will plausibly be preferred when thinking about Dave (in order to make the question sensible), and insofar as this conception will plausibly be preserved in the next question about Kate.

Yet again we found no significant order effects on people's attributions of knowledge³⁸ or belief.³⁹ We found that 83% of participants were happy to ascribe belief to Dave despite his being asleep. And most crucially we continued to find that the majority response (47%) was to attribute both knowledge and belief to Kate:



There were no significant order effects on people's attributions of belief, df=2, X^2 =2.215, p=.137.

-

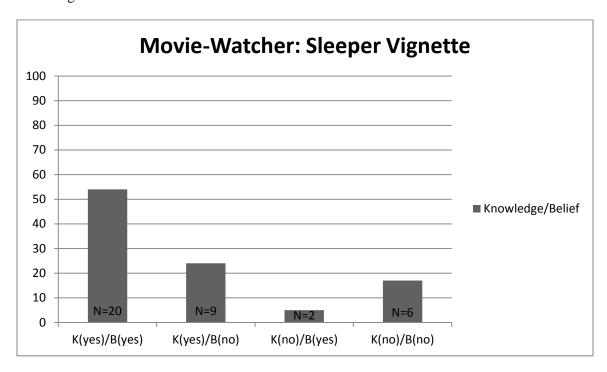
³⁸ There were no significant order effects on people's attributions of knowledge, df=2, X^2 =.295, p=.587.

Among the subgroup that ascribed knowledge, the majority (72%) also ascribed belief. ⁴⁰ And among the subgroup that ascribed belief, the majority (81%) also ascribed knowledge. 41

We then tried the very same sort of manipulation with Myers-Schulz and Schwitzgebel's case of the freaked-out movie-watcher, adding a character—Alex—into the story who had fallen asleep during the movie (V4). We then (while retaining a within-subjects design, and counterbalancing the order of the knowledge and belief probes) asked people about what Alex believed before asking about what Jamie believed:

Did Alex (despite being asleep) believe that only water would come out of the sink faucet?

Yet again we found no significant order effects on people's attributions of knowledge⁴² or belief. 43 We found that 89% of participants were happy to ascribe belief to Alex despite his being asleep. And most crucially we continued to find that the majority response (54%) was to attribute both knowledge and belief to Jamie:



Among the subgroup that ascribed knowledge, the majority (69%) also ascribed belief. 44 And among the subgroup that ascribed belief, the majority (90%) also ascribed knowledge.⁴⁵

⁴⁰ Among those who ascribed knowledge, the proportion of those who also ascribed belief (25) is greater than the proportion of those who denied belief (10), p<.001, binomial test proportion=.29.

41 Among those who ascribed belief, the proportion of those who also ascribed knowledge (25) is greater than the

proportion of those who denied knowledge (6), p<.001, binomial test proportion=.19. ⁴² There were no significant order effects on people's attributions of knowledge, df=2, X^2 =.069, p=.763.

⁴³ There were no significant order effects on people's attributions of belief, df=2, X^2 =2.042, p=.135.

⁴⁴ Among those who ascribed knowledge, the proportion of those who also ascribed belief (20) is greater than the proportion of those who denied belief (9), p<.001, binomial test proportion=.31.

45 Among those who ascribed knowledge, the proportion of those who also ascribed belief (20) is greater than the

proportion of those who denied belief (2), p<.001, binomial test proportion=.09.

We take the uniformity of these results to buttress the claim that both of our revisions succeeded in eliciting the dispositional reading of 'believe,' and thereby to confirm our prediction (§1) that ordinary speakers will say that knowledge entails dispositional belief. As one final buttressing design, we went back to *Unconfident examinee* (V1) and simply replaced the belief probe with:

Is Kate disposed to believe that Queen Elizabeth died in 1603?

We found—in conformity with our previous results—that 55% ascribed both knowledge and belief, that among the subgroup that ascribed knowledge 69% also ascribed belief, and that among the subgroup that ascribed belief 77% also ascribed knowledge. So we conclude that, at least for two out of the three studies for which Myers-Schulz and Schwitzgebel found evidence against the entailment thesis, specifically eliciting the relevant dispositional reading of 'believe' reverses their results, and thereby provides *confirmation* for the entailment thesis properly understood. We thereby claim confirmation of our own empirical predictions from §1.3.

Just to clarify our minimal commitments with respect to eliciting readings of 'believe,' we are committed to the following:

- There are multiple possible meanings for 'believe'
- Myers-Schulz and Schwitzgebel's studies did not uniformly elicit the epistemically relevant reading
- The techniques we used to re-do their studies did more to uniformly elicit the epistemically relevant reading⁴⁷

We further suspect that Myers-Schulz and Schwitzgebel's unconfident examinee and freaked-out movie-watcher studies tended to elicit the *occurrent* reading, but strictly speaking that is inessential for the current argument. We also would characterize the epistemically relevant reading as the *dispositional* reading, but strictly speaking that is inessential too. Our first claim—that there are multiple possible meanings for 'believe'—is buttressed by the fact that we see very different data on simple belief probes between the vignettes without the sleeper (V1, V3) and their counterpart vignettes with the sleeper (V2, V4). Our further two claims—that our studies are doing more to uniformly elicit the epistemically relevant reading—then fall out of the guiding conception that the epistemically relevant reading of 'believe' is in terms of a notion applicable even to those who are asleep (§1.2).

3.2 Confirming the fourth worry: neutralizing moral valence

Myers-Schulz and Schwitzgebel also claimed to find evidence against the entailment thesis from their case of the prejudiced professor. We have not yet addressed this last case. First of all, we think that this case is very different from that of the unconfident examinee or freaked-out movie-watcher cases (both of which we think of as cases in which the subject clearly has dispositional belief). The prejudiced professor—Juliet—is not a character who has a single standing belief that is merely temporarily masked by extreme circumstances. (For that reason we do not think that controlling for a disposition reading of 'believe' is relevant.) Rather Juliet exhibits an enduring conflict between standing beliefs. She is an internally deeply conflicted character for whom ascription of any relevant standing beliefs or knowledge

⁴⁶ We additionally take the difference between the results Myers-Schulz and Schwitzgebel report for their belief probe, and the results we report for the same probe in our revised sleeper vignettes, as some confirmation of our claim (§1.2) that 'believe' has both an occurrent and a dispositional sense.

⁴⁷ Strictly speaking this third claim entails the previous second claim, but we find it useful to separate these claims since they might be contested separately. After all, our second claim marks a worry with Myers-Schulz and Schwitzgebel's studies, while our third claim mainly concerns the success of our re-designs.

is tricky. So our guiding intuition is that—far from being a clear case of knowledge without belief—the case is simply unclear in all directions.⁴⁸

Rather we think—in accord with our fourth worry—that the case of the prejudiced professor involves the complicating matter of moral valence. Juliet is said to have prejudicial tendencies, of a sort that will plausibly elicit knee-jerk moral condemnation from most participants. Could it be that, while participants are split on whether Juliet believes, their tendency to say that Juliet knows can be explained away (perhaps as a matter of cognitive bias) from the moral valence of the case?

Indeed, recent work by Beebee and Buckwalter (2010) has shown that knowledge ascriptions can be influenced by moral considerations. In particular, they found that knowledge ascriptions are increased when the target agent's behavior is negatively valenced (as Juliet's is) and decreased when the target agent's behavior is neutral or positively valenced. So there is already an empirical basis for suspecting that the negative valence attached to Juliet's prejudice is elevating ascriptions of knowledge. Our question is whether Myers-Schulz and Schwitzgebel's results can be replicated in a morally neutral counterpart case.

Accordingly we replaced Myers-Schulz and Schwitzgebel's vignette (V5) with a vignette about a merely absentminded professor, who in reflective thought thinks it is a Saturday, but in action behaves as if it were a Wednesday (V6). Our guiding thought was that absentmindedness is a relatively morally neutral status, so that our revised case would test whether internally conflicted subjects like Juliet elicit intuitions of knowledge without belief, without any potentially distorting effects of moral valence.

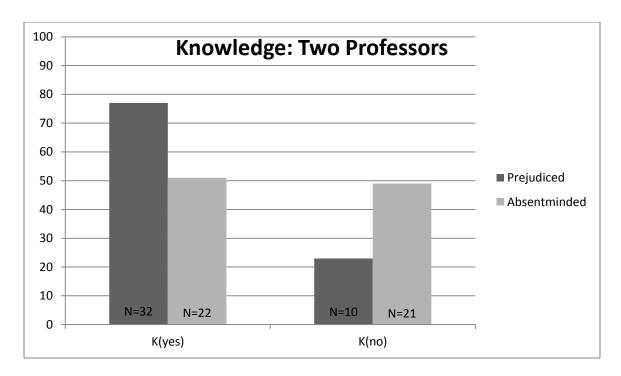
We retained our usual within-subjects design, counterbalancing the order of questions. We also re-did Myers-Schulz and Schwitzgebel's prejudiced professor case using a within-subjects design, just to ensure that we were comparing the prejudiced and the absentminded professor cases under comparable designs. We will report the results side-by-side.

No order effects were found with either the prejudiced or absentminded professor vignettes, for either knowledge⁴⁹ or belief ⁵⁰ ascriptions. But we did find—in accord with our concern about moral valence, and in accord with Beebe and Buckwalter's (2010) findings—that assignment to the prejudiced versus absentminded professor cases had an effect on knowledge attributions:⁵¹

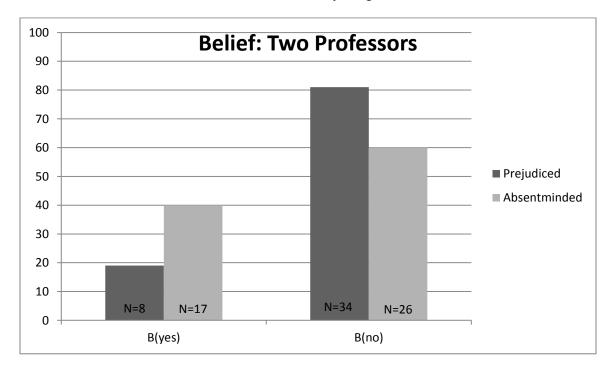
⁴⁸ In the useful terms of Gendler (2008), Juliet might be said to *believe* that student athletes are just as capable of other students, but not to alieve that same proposition, where "alief" is associated with more instinctual "animalistic" type 1 cognitive systems, and "belief" is associated with more reflective "rationalistic" type 2 cognitive systems. But of course Gendler is proposing her alief/belief distinction as a useful revisionary taxonomy, precisely to overcome potentially conflationary aspects in the folk notion of belief. So we suspect that it may be somewhat indeterminate whether or not the folk notion of belief applies to a conflicted character like Juliet.

⁴⁹ In the prejudiced professor case, there were no significant order effects on people's attributions of knowledge, df=2, X²=.260, p=.610. In the absentminded professor case there were also no significant order effects on people's attributions of knowledge, df=2, X^2 =.568, p=.451.

⁵⁰ In the prejudiced professor case, there were no significant order effects on people's attributions of belief, df=2, $X^2=2.898$, p=.090. In the absentminded professor case there were also no significant order effects on people's attributions of belief, df=2, X^2 =1.608, p=.205. ⁵¹ df=2, X^2 =5.744, p=.017.



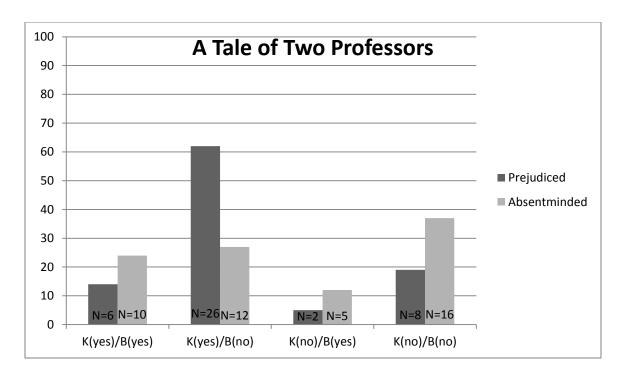
We also found that belief attributions were affected by assignment to either condition:⁵²



So it seems—perhaps unsurprisingly, but not in a way that had hitherto been documented—that there is also a "Knobe effect" for belief ascription. The overall response patterns were:

21

 $^{^{52}}$ df=2, X^2 =4.295, p=.038.



So we conclude that cases of internally conflicted subjects are difficult cases for the ascription of folk attitudes, and that it is only in the negatively morally valenced versions of such cases like that of the prejudiced professor that we see a dominant tendency to ascribe knowledge without belief. In the absentminded professor case the responses are "all over the map," with the majority (a bare 37%) denying both knowledge and belief. This fits our guiding intuition that the case—far from being an intuitive case of knowledge without belief—is really, when considered in the standard terms of folk psychology, an unclear case in all directions.

What one makes of these results will depend on how one thinks about the Knobe effect generally. If the Knobe effect is due to a cognitive bias, then one should conclude that Myers-Schulz and Schwitzgebel's results are merely due to an independent bias, and that removing such a bias restores people's tendency to intuit in accord with the entailment thesis (thus our talk about "removing biases" in the introduction). For instance, at least one leading view—associated with Alicke—is that the Knobe effect is due to cognitive distortions induced by an overriding desire to blame. As Alicke and Rose (2010: 330) summarize: "The culpable control model, therefore, does not require the assumption that concepts such as intentionality, causation, and foresight are suffused with moral considerations. Rather, the influences of these evaluations can be explained in terms of the desire to blame an agent whose actions arouse strong disapproval."

If the Knobe effect is not due to a cognitive bias then it may still be that Myers-Schulz and Schwitzgebel have found an intuitive counterexample to the claim that knowledge entails belief. Indeed their counterexample would be a new and fascinating counterexample quite distinct from cases like *The case of the unconfident examinee*, which (we have argued: §1, §3.1) clearly involve cases of dispositional belief. It would be a counterexample that combined conflicts between enduring beliefs with moral valence. In that eventuality our contribution would be one of disentangling the cases, and showing that Myers-Schulz and Schwitzgebel did not merely "empirically verify Radford," but actually achieved something far more significant than they advertise, in providing and empirically confirming an entirely new style of counterexample to the entailment thesis!

But for the time being—given that it remains an open question as to how to understand the Knobe effect⁵³—we can only conclude that Myers-Schulz and Schwitzgebel have not yet established a counterexample to the entailment thesis. Orthodoxy should still stand. (The burden of proof is of course on those who would topple orthodoxy.) But its standing becomes—in surprising and interesting ways that we credit Myers-Schulz and Schwitzgebel for revealing—dependent on the open question of how best to understand the Knobe effect.

3.3 Speculations on the fifth worry: explaining the diversity

So far we take ourselves to have explained away each of the three studies Myers-Schulz and Schwitzgebel offered in which people seemed to have intuitions that conflicted with the entailment thesis. With the unconfident examinee and the freaked-out movie-watcher, we were able to show that specifically eliciting the dispositional reading of 'believe' reversed their results, and with the prejudiced professor we were able to show that neutralizing moral valence revealed a case in which intuitions were simply unclear in all respects.

It remains to consider our fifth and final worry about Myers-Schulz and Schwitzgebel's studies, which is that they found quite diverse results between their five studies, without offering any explanation for the diversity of their results. (Such diverse results seem to fit neither orthodoxy nor Myers-Schulz and Schwitzgebel's preferred replacement: §2.1.) Though we will not attempt further empirical confirmation of this point, we would speculate that the reason why their absent-minded driver and self-deceived husband cases (V7, V8) did *not* show a statistically significant tendency towards ascriptions of knowledge without belief, is that these cases did not do nearly as much as trigger the problems we've tried to expose above.

The main difference we see between the unconfident examinee and freaked-out movie-watcher cases (V1, V3) on the one hand, and the absent-minded driver case (V7) on the other, is whether the vignette focuses on the phenomenology of the subject. In both the unconfident examinee and freaked-out movie-watcher vignettes, participants are given *detailed and vivid* descriptions of how the subject feels. (Indeed the relevant unconfidence and freaked-outedness are highlighted in the very titles of these vignettes.) But in the absent-minded driver we don't get this detailed and vivid window into the mental life of the driver. We are only told that Ben is "enthusiastically tapping his fingers to the music, not paying much attention to where he is going." Little is said of Ben's occurrent phenomenology (as befits a case of "absent-mindedness"). We speculate that this comparative difference in the detailed and vivid phenomenological descriptions is making a difference in whether participants are naturally drawn to the irrelevant occurrent conception of belief (since this is the phenomenologically loaded conception). If so then we have an explanation—one that crucially involves the occurrent/dispositional distinction—for why the data patterns came out differently for these cases.

With the self-deceived husband case (V8) we see a case that is more analogous to the prejudiced professor case (V5), in terms of featuring a deeply internally conflicted subject (as per the idea of self-deception highlighted in the title of the one vignette). We speculate that the main difference is that being prejudiced carries clearly negative moral valence, while the moral status of the self-deceived husband is more complex and less clearly negative (there is something both sad and noble about Tim's clinging to the belief of his wife's faithfulness; it is natural to have mixed feelings about Tim). We thus see the self-deceived husband case as akin to our absent-minded professor case (V6). So we speculate that the self-deceived husband case is generating different results from the prejudiced professor case insofar as the crucial driver of the prejudiced professor case—negative moral valence—goes missing.

-

⁵³ For further discussion see Knobe 2010 and associated commentaries.

This section has been speculative, but in a way that fits the pattern of empirical results reported above. Of course further empirical work may suggest a still different overall picture. There is far more work to be done, and we mainly hope to inspire others to explore these issues further.

Putting all of this together, we have offered a new philosophical reply to *Unconfident examinee*, alongside new empirical data which we take to reverse Myers-Schulz and Schwitzgebel's case against the orthodox view that knowledge entails belief. We thus conclude—shocking as it may sound for work in experimental philosophy—that the orthodox armchair view withstands empirical scrutiny.⁵⁴

Appendix: Vignettes

V1 Unconfident Examinee

Kate spent many hours studying for her history exam. She's now in class taking the exam. Everything's going quite well, until she comes to the final question. It reads, "What year did Queen Elizabeth die?" Kate had reviewed this date many times. She had even recited the date to a friend just a few hours earlier. So, when Kate sees that this is the last question, she feels relieved. She confidently looks down at the blank space, waiting to recollect the answer. But before she can remember it, the teacher interrupts and announces, "Alright, the class session is almost over. You have one more minute to finalize your answers." Kate's demeanor suddenly changes. She glances up at the clock, now flustered and worried. "Oh, no. I can't perform well under this kind of pressure." Her grip tightens around her pencil. She strains to recall the answer, but nothing comes to her. She quickly loses confidence. "I suppose I'll just have to guess the answer," she says to herself. With a sigh of disappointment, she decides to write "1603" into the blank space. This was, in fact, the correct answer.

V2 Unconfident Examinee with Sleeper

Kate and Dave both spent many hours studying for their history exam together. Kate is now in the class taking the exam, while poor Dave has slept through his alarm clock and is still sound asleep in bed. Kate is doing quite well on the exam, until she comes to the final question. It reads, "What year did Queen Elizabeth die?" Kate and Dave had reviewed this date many times. They had even recited the date to each other at the end of their last study session. So, when Kate sees that this is the last question, she feels relieved. She confidently looks down at the blank space, waiting to recollect the answer. But before she can remember it, the teacher interrupts and announces, "Alright, the class session is almost over. You have one more minute to finalize your answers." Kate's demeanor suddenly changes. She glances up at the clock, now flustered and worried. "Oh, no. I can't perform well under this kind of pressure." Her grip tightens around her pencil. She strains to recall the answer, but nothing comes to her. She quickly loses confidence. "I suppose I'll just have to guess the answer," she says to herself. With a sigh of disappointment, she decides to write "1603" into the blank space. This was, in fact, the correct answer.

V3 Freaked-Out Movie-Watcher

Susan loves to watch old horror films. She finally convinces her friend Jamie to watch one with her. It's an old horror film that Susan actually considers to be quite funny, due to its unrealistic plot. The film begins with a group of astronauts who discover alien life on another planet. The aliens look somewhat like bumblebees, but they are dark-green and about two feet in length. The astronauts capture one of these creatures and bring it back to Earth. Once they have it on Earth, it manages to escape and starts laying numerous eggs. The eggs need water to hatch, so the creature lays the eggs in sink faucets. Thus, whenever people turn on their sink faucet, hundreds of newly hatched alien creatures fly out and begin to attack them.

During one of these attack scenes, Susan notices that Jamie is a bit tense. Susan remarks, "This isn't bothering you, is it? Come on, you should be laughing at this movie. Look how unrealistic it is." Jamie responds, "Yes, of course it's unrealistic. But it's still scary. I just don't like these types of movies.

⁵⁴ Thanks to Wesley Buckwalter, Joshua Knobe, Blake Myers-Schulz, Eric Schwitzgebel, and Justin Sytsma.

They frighten me. Can't we just watch something else?" "Well, I suppose," Susan says. Susan then turns off the movie, and they quickly get ready for a second trip to the movie store.

On the way out, Susan stops. "Hold on for a second. I'm thirsty. Let me grab a glass of water." Susan walks over and begins to turn on the sink faucet. Suddenly, Jamie shouts, "No! Don't do it!" The words come out of Jamie's mouth before she even has time to consider what she's saying. Jamie then looks over and sees that it's only water coming out of the faucet.

V4 Freaked-Out Movie-Watcher with Sleeper

Susan loves to watch old horror films. She finally convinces her friends Jamie and Alex to watch one with her. It's an old horror film that Susan actually considers to be quite funny, due to its unrealistic plot. The film begins with a group of astronauts who discover alien life on another planet. The aliens look somewhat like bumblebees, but they are dark-green and about two feet in length. The astronauts capture one of these creatures and bring it back to Earth. Once they have it on Earth, it manages to escape and starts laying numerous eggs. The eggs need water to hatch, so the creature lays the eggs in sink faucets. Thus, whenever people turn on their sink faucet, hundreds of newly hatched alien creatures fly out and begin to attack them.

During one of these attack scenes, Susan notices that Alex has fallen soundly asleep, while Jamie is a bit tense. Susan ignores Alex's snoring and says to Jamie: "This movie isn't bothering you, is it? Come on, you should be laughing at this movie. Look how unrealistic it is." Jamie responds, "Yes, of course it's unrealistic. But it's still scary. I just don't like these types of movies. They frighten me. Can't we just watch something else?" "Well, I suppose," Susan says. Susan then turns off the movie, and they quickly get ready to leave Alex snoring on the couch and take a second trip to the movie store.

On the way out, Susan stops. "Hold on for a second. I'm thirsty. Let me grab a glass of water." Susan walks over and begins to turn on the sink faucet. Suddenly, Jamie shouts, "No! Don't do it!" The words come out of Jamie's mouth before she even has time to consider what she's saying. Jamie then looks over and sees that it's only water coming out of the faucet.

V5 Prejudiced Professor

Juliet is a university professor. Unfortunately, she is also prejudiced against student athletes. In her classes, she calls more often on non-athletes than athletes, and she interprets the comments of the former more charitably. When two soccer players, Brett and Bernard, come to visit her in office hours, she treats them patronizingly, explaining the basic concepts of the course in a very rudimentary manner, failing to recognize the sophistication and intelligence behind their questions. They leave, and shortly after, two students with no involvement in school sports enter. Juliet immediately launches into a high-level discussion, generously assuming the students' command of the elementary material. When Bernard writes the best essay in the course, revealing the intelligence that a neutral observer would have recognized in his previous remarks, Juliet is surprised. All of this is typical of her.

However, Juliet also repudiates all forms of prejudice. She openly affirms that students involved in athletics are just as capable as non-athletes. In fact, she has it on excellent authority that this is the case: Her chair just completed a study showing that the two groups perform equally well in their philosophy classes. Intrigued by this study, Juliet even reviews her own records and finds that, on average, the athletic students had actually performed better than the other students. But, in spite of all this, Juliet's prejudice remains. She continues to treat her athletic students as if they are less intelligent than her other students.

V6 Absentminded Professor

Juliet is a university professor. Unfortunately, she is quite absentminded. She has been walking around all morning thinking that it is "a beautiful Saturday" when in fact it is a Wednesday. Since she thinks it is Saturday, she has been considering driving to the countryside for a scenic picnic as she usually does on Saturday's. Indeed, she has it on "excellent authority" that it is a Saturday: Her normally reliable

cellphone, which Juliet had accidentally run through the washing machine last night, now lists the date as "Saturday September 9th."

But in spite of the fact that some part of her mind thinks that today is a Saturday, Juliet in fact behaves just as she does every Wednesday. She drives to school as usual and shows up to teach her Wednesday class on time. She holds her Wednesday office hours as scheduled, and then meets up with a colleague for their usual Wednesday afternoon coffee and brainstorming session. She never reconciles her idea that today is Saturday with any of these actions. This sort of behavior is typical of absentminded Juliet.

V7 Self-Deceived Husband

Tim's wife Diane is cheating on him. For two years, Diane has been conducting a romantic affair with Mark, who is a colleague of hers at work. Over the past two years, Tim has seen frequent clues that Diane is cheating: unexpected credit card charges, late arrivals from work with weak and flustered explanations as to why, unexplained mysterious phone calls, etc. Diane even occasionally calls Tim "Mark", and once Tim overheard her saying "I love you, Mark" on the telephone when Diane assumed Tim was not in the house. One night several months ago, Diane even confessed to him explicitly, saying anxiously in a quiet moment in bed, "Tim, you know that I have fallen in love with another man and have been cheating on you for a couple of years". Tim loudly insisted that she was joking, just trying to get his goat because she was mad with him about some out-of-town travel he was doing – and Diane replied that, yes, she was of course just joking.

Despite all this evidence, Tim vehemently insists that his marriage is in good shape and that Diane would never even think of cheating on him. Perhaps, indeed, he says such things a little *too* vehemently. When Dan, a friend of Tim's, gently points out to Tim some of the evidence of Diane's affair, Tim dismisses Dan's remarks as utter nonsense, saying to himself, "Dan is probably just jealous and wishes that his own marriage was as solid as Diane's and mine". When a woman whom Tim finds attractive starts flirting with him at work, Tim brushes her off, saying to himself that he could never do anything that might threaten his marriage. At the same time, however, when Diane comes home late, Tim finds himself much more anxious and bothered about it than he ever used to be, though he can't quite put his finger on why. When he answers the phone and finds no one there, he sometimes finds himself wondering "could it be a lover of Diane's?" and then, very quickly after that, "Ridiculous! Ridiculous! She would never cheat!" When he sees a credit card charge for an 8:00 pm dinner at a romantic restaurant, he finds himself with a visual image of Diane having a romantic dinner with a stranger – an image which he rejects as a horrible fantasy, but that he can't quite put out of mind.

V8 Absent-Minded Driver

Ben receives an email informing him of a bridge closure on his normal route to work. He becomes mildly annoyed and says to himself, "Now I'll have to turn on Russell Street and go all the way down to Langdon Avenue." So, the next morning, Ben wakes up early and quickly gets ready for work. He makes it out of the house with plenty of time to make the drive. Pleased with the success of his early departure, he decides to listen to one of his favorite albums and enjoy the long drive. By the time Ben is approaching Russell Street, where he should turn, he is enthusiastically tapping his fingers to the music, not paying much attention to where he is going, and he drives right past Russell Street, continuing on his normal route to work. Thus it's only a matter of time before Ben will reach the closed bridge and have to drive all the way back to Russell Street. Nevertheless, Ben just keeps on tapping his fingers to the music and continues to drive towards the closed bridge.

References

Alicke, Mark 1992. Culpable Causation. *Journal of Personality and Social Psychology* 63: 368-78. Alicke, Mark and David Rose 2010. Culpable Control or Moral Concepts? *Behavioral and Brian Sciences* 33: 330-31.

Armstrong, David 1969. Does Knowledge Entail Belief? *Proceedings of the Aristotelian Society* 70: 21-36

Audi, Robert 1994. Dispositional Beliefs and Dispositions to Believe. Nous 28: 419-34.

Black, Carolyn 1971. Knowledge without Belief. Analysis 31: 152-58.

Beebee, James and Wesley Buckwalter 2010. The Epistemic Side-Effect Effect. *Mind & Language* 25: 474-98.

Buckwalter, Wesley manuscript. Factive Verbs and Protagonist Projection.

Campbell, C. A. 1967. Towards a Definition of Belief. The Philosophical Quarterly 17: 204-20.

Cohen, L. Jonathan 1966. More about Knowing and Feeling Sure. Analysis 27: 11-16.

— 1989. Belief and Acceptance. *Mind* 98: 367-89.

Dartnall, Terry 1986. Radford Revisited. Philosophical Quarterly 36: 395-98

Gendler, Tamar 2008. Alief and Belief. Journal of Philosophy 105: 634-63.

Gettier, Edmund 1963. Is Justified True Belief Knowledge? Analysis 23: 121-23.

Goodman, Nelson 1955. Fact, Fiction, and Forecast. Harvard University Press.

Harrison, Jonathan 1963. Does Knowing Imply Believing? *Philosophical Quarterly* 13: 322-32.

Horgan, Terence and John Tienson 2002. The Intentionality of Phenomenology and the Phenomenology of Intentionality. *Philosophy of Mind: Classical and Contemporary Readings*, ed. Chalmers: 520-33. Oxford University Press.

Johnston, Mark 1992. How to Speak of the Colors. *Philosophical Studies* 68: 221-63.

Jones, O. R. 1971. Knowing and Guessing: By Examples. Analysis 32: 19-23.

Klein, Peter 1999. Human Knowledge and the Infinite Regress of Reasoning. *Philosophical Perspectives* 13: 297-325.

Knobe, Joshua 2003. Intentional Action and Side-Effects in Ordinary Language. Analysis 63: 190-93.

— 2010. Person as Scientist, Person as Moralist. Behavioral and Brian Sciences 33: 315-29.

Knobe, Joshua and Shaun Nichols 2008. An Experimental Philosophy Manifesto. *Experimental Philosophy*, eds. Knobe and Nichols: 3-14. Oxford University Press.

Lehrer, Keith 1968. Belief and Knowledge. Philosophical Review 77: 491-99.

— 1974. *Knowledge*. Clarendon Press.

Lewis, David 1996. Elusive Knowledge. Australasian Journal of Philosophy 74: 549-67.

Ludlow, Peter 2008. Cheap Contextualism. Philosophical Issues 18: 104-29.

Lycan, William 1988. Judgment and Justification. Oxford University Press.

Mannison, D. S. 1976. "Inexplicable Knowledge" does not Require Belief. *Philosophical Quarterly* 26: 139-48.

Margolis, Joseph 1973. Knowledge and Existence. Oxford University Press.

Myers-Schulz Blake and Eric Schwitzgebel *forthcoming*. Knowing that P without Believing that P. *Nous*. Price, H. H. 1969. *Belief*. Allen & Unwin.

Prichard, H. A. 1950. Knowledge and Perception. Oxford University Press.

Radford 1966. Knowledge—By Examples. Analysis 27: 1-11

- 1970a. Does Unwitting Knowledge Entail Unconscious Belief? *Analysis* 30: 103-7.
- 1970b. Analyzing "Know(s) That," *Philosophical Quarterly* 20: 222-29.
- 1972. On Sticking to What I Don't Believe to Be the Case. *Analysis* 32: 170-73.
- 1988. Radford Revisiting. *Philosophical Quarterly* 38: 496-99.
- 1990. Belief, Acceptance, and Knowledge. *Mind* 99: 609-17.

Ryle, Gilbert 1949. The Concept of Mind. Hutchinson.

Schaffer, Jonathan 2005. Contrastive Knowledge. Oxford Studies in Epistemology 1: 235-71.

Schope, Robert 2002. Conditions and Analyses of Knowing. *The Oxford Handbook of Epistemology*, ed. Moser: 25-70. Oxford University Press.

Schwitzgebel, Eric 2001. In-Between Believing. Philosophical Quarterly 51: 76-82.

- 2002. A Phenomenal, Dispositional Account of Belief. *Nous* 36: 249-75.
- 2010a. Acting Contrary to our Professed Beliefs, or the Gulf Between Occurrent Judgment and Dispositional Belief. *Pacific Philosophical Quarterly* 91: 531-53.

— 2010b. Belief. Stanford Encyclopedia of Philosophy.

Sorensen, Roy 1982. Knowing, Believing, and Guessing. Analysis 42: 212-13.

Steup, Matthias 2006. The Analysis of Knowledge. Stanford Encyclopedia of Philosophy.

Sytsma, Justin and Jonathan Livengood 2011. A New Perspective Concerning Experiments on Semantic Intuitions. *Australasian Journal of Philosophy* 89: 315-32.

Vendler, Zeno 1972. Res Cogitans: An Essay in Rational Psychology. Cornell University Press.

— 1978. Escaping from the Cave: A Reply to Dunn and Suter. *Canadian Journal of Philosophy* 8: 79-87.

Williamson, Timothy 2000. Knowledge and its Limits. Oxford University Press.

Woozley, A. D. 1948. Dispositions. *Mind* 57: 350-53.

— 1952. Knowing and Not Knowing. *Proceedings of the Aristotelian Society* 53: 151-72.