# The problem of ESEE knowledge<sup>*</sup>

JOHN TURRI
john.turri@gmail.com

*Abstract*: Traditionally it has been thought that the moral valence of a proposition is, strictly speaking, irrelevant to whether someone knows that the proposition is true, and thus irrelevant to the truth-value of a knowledge ascription. On this view, it's no easier to know, for example, that a bad thing will happen than that a good thing will happen (other things being equal). But a series of very surprising recent experiments suggest that this is actually not how we view knowledge. On the contrary, people are much more willing to ascribe knowledge of a bad outcome. This is known as the epistemic side-effect effect (ESEE) and is a specific instance of a widely documented phenomenon, the side-effect effect (a.k.a. "the Knobe effect"), which is the most famous finding in experimental philosophy. In this paper, I report a new series of five experiments on ESEE and in the process accomplish three things. First, I confirm earlier findings on the effect. Second, I show that the effect is virtually unlimited. Third, I introduce a new technique for detecting the effect, which potentially enhances its theoretical significance. In particular, my findings make it more likely, though they do not entail, that the effect genuinely reflects the way we think about and ascribe knowledge, rather than being the result of a performance error.

## 1. Introduction

Most theorists think that whether a belief counts as knowledge is unaffected by the desirability of the belief or its content. If we're wondering whether you know that a certain event will occur, we'll attend to whether you believe that the event will occur, whether it's true that the event will occur, whether you have any evid-

---

<sup>*</sup> This is the penultimate version of a paper forthcoming in *Ergo*. Please cite the final, published version if possible.

ence that it will occur, and perhaps also whether your evidence only "luckily" led you to believe the truth. It doesn't matter whether the event in question is good or bad. That is, other things being equal, the event's valence doesn't affect whether you know that it will occur. Or so mainstream theorizing about knowledge would have have us believe (e.g. Steup 1996, Lehrer 2000, Feldman 2003, Fumerton 2006, BonJour 2009).

But important recent experimental work suggests that mainstream opinion is simply wrong about this (Beebe and Buckwalter 2010, Beebe and Jensen 2012, Beebe and Shea forthcoming). The valence of an outcome apparently does influence whether we ascribe knowledge. For example, people are more willing to ascribe knowledge to a CEO if he thinks that his company's policies will *harm* rather than *help* the environment. This is known as *the epistemic side-effect effect* (ESEE; pronounced 'easy'). It is a specific instance of a widely documented phenomenon, *the side-effect effect* (a.k.a. "the Knobe effect"), which is the most famous finding in experimental philosophy. ESEE is an important and surprising discovery about how we think of knowledge, which should be of interest to philosophers and cognitive scientists more generally. It is my focus here.

Contemporary epistemologists have taken it for granted that patterns in ordinary usage should, or at least can, constrain substantive epistemological theorizing. It is widely assumed that theorists should respect patterns of competent and literal knowledge attribution. Other things being equal, we prefer an epistemological theory to the extent that it counts ordinary knowledge attributions as competent and literal (more on this in section 2.2). In this paper, I investigate whether we should think that ESEE reflects such a pattern and, in turn, whether it ought to inform epistemological theory. Although further work ultimately remains to be done, I report some good initial evidence that ESEE does reflect a pattern worthy of theoretical respect.

Here is the paper's plan. Section 2 situates the present discussion within a

broader research program; it reviews previous work on side-effect effects and outlines competing approaches to explaining them. Section 3 poses two questions that set the stage for further experimental inquiry into ESEE. Section 4 reports four experiments which collectively suggest that, when observed using previous methods, ESEE is too broad and easily activated for it to reveal how we think about knowledge. Section 5 suggests an alternative approach for detecting ESEE. Section 6 enacts this alternative and reports an experiment that provides the best evidence thus far that ESEE reflects a competent pattern of knowledge ascription. Sections 7 sets an agenda for further research into this topic.

## 2. Knowledge can be ESEE

There has been a spate of valuable recent work on ordinary knowledge ascriptions (e.g. Weinberg et al. 2001, Swain and Weinberg 2008, Nagel 2008, May et al. 2010, Feltz and Zarpentine 2010, Starmans and Friedman 2012, Schaffer and Knobe 2012, Sripada and Stanley forthcoming, Pinillos forthcoming). These studies suggest that some surprising factors affect whether laypeople attribute knowledge (for an overview, see Buckwalter 2012). Here I will focus on one set of studies in particular. These studies take inspiration from Joshua Knobe's pioneering work on how people's evaluative judgments affect whether they describe outcomes as *intentional* (Knobe 2003a, 2003b, 2004, 2010). This section briefly reviews relevant prior work.

### 2.1. SEE, this is easy

Suppose that an agent embarks on a course of action aimed at a primary outcome (a central effect). And suppose that the agent also anticipates but does not desire that her conduct will have a certain side effect. Knobe found that if the side effect is bad or undesirable, then people are much more likely to describe the agent as

having *intentionally* brought it about. For example, to use Knobe's now famous case, suppose that a CEO decides to start a new program with the primary aim being to increase profits. And suppose that the CEO also anticipates but does not desire that the program will, as a side effect, affect the environment. Did the CEO intentionally bring about the environmental side effect? People are much more likely to say that he did if the side effect is *harmful* rather than *helpful*. This is known as *the side-effect effect* (SEE).

A side effect's valence influences ascriptions of intentionality, and it has also been shown to affect ascriptions of other practical attitudes, such as *deciding*, *being in favor of*, and *desiring*, among others (Pettit and Knobe 2009, Tannenbaum et al. 2009), as well as *act individuation* (Ulatowski 2009). This is true whether the side effect is negatively valenced because it violates a norm of morality, prudence, aesthetics, or etiquette. The effect appears to be present regardless of gender, age or native language. The effect is very robust.

An important question is what explains SEE. Knobe groups the competing accounts into *competence explanations* and *distortion explanations*, and this taxonomy has taken hold in the literature (e.g. Buckwalter 2013). Competence explanations attribute the effect to the fact that 'moral considerations actually figure in the competencies people use to' attribute psychological states (Knobe 2010: 320). Knobe favors this sort of account, though he does not claim that it has any immediate implications for a substantive theory of what psychological states are, or for a semantic theory of psychological predicates (e.g. Knobe and Burra 2006: 332–3, 338). Instead, Knobe claims only that the effect reveals something deep and important about the way we ordinarily think about psychological states, about our underlying competence in applying such concepts.

Distortion explanations deny that moral considerations figure into the underlying competence, and instead attribute the effect to other factors that distort the competence's operation. They claim that the effect is instead due to performance

error (Nadelhoffer 2006, Malle 2006: 103–4, Alicke and Rose 2010), or to pragmatic concerns that don't reveal anything about how people think about psychological states *per se* (Adams and Steadman 2004), or to features of the materials used in the experiments (Guglielmo and Malle 2010).

Knobe gently sets aside whether there is a semantic or metaphysical upshot of his conceptual competence model. But due to certain preoccupations and methodological commitments in contemporary epistemology, when we turn to whether the side-effect effect shows up for knowledge ascriptions too, it's not so simple to maintain the division between people's underlying competence in classifying things in a certain category, on the one hand, and the nature of things that populate that category, on the other.[1]

Epistemologists have taken it for granted that patterns in ordinary usage should, or at least can, constrain substantive theorizing about the nature of knowledge, or the (abstract) concept of knowledge, or the meaning of 'knows' (Austin 1956, Cohen 1999, Rysiew 2001, DeRose 2009, Hawthorne 2004, Stanley 2005, Fantl and McGrath 2009). Of course, not just any pattern is ripe for guiding theory. But it is widely assumed that theorists should respect patterns of competent and literal knowledge ascription. There will always be some interplay between one's theory and what one is willing to consider a pattern of competent and literal knowledge ascription (e.g. Davis 2007, Bach 2008, Turri forthcoming b). But it's not "anything goes." And a theory which implies that much of our ordinary practice of knowledge ascription is either incompetent or non-literal has, as they say, "some explaining to do."

Setting aside contemporary epistemologists' methodological predilections, there is another reason to take competent and literal application of concepts seriously. Over the past forty years, philosophers of language and mind have de-

---

[1]  Buckwalter (2013) does an especially good job of explaining and properly emphasizing the significance of these methodological points.

veloped *externalist* theories of semantic meaning and mental content (Kripke 1972, Putnam 1975, Burge 1979, Burge 1986). Though not uncontroversial, these externalist theories are widely accepted. If, as Donald Davidson (1983: 146) puts it, 'Belief is in its nature veridical,' or as George Santayana (1923: 9) wrote, 'Intelligence is by its nature veridical,' or as the 11th century Indian philosopher Sriharsa thought, 'Cognition is ordinarily by nature true or veridical' (Phillips 2011), then we should expect competent and literal applications of concepts to be largely accurate. So despite Knobe's attempt to downplay it, a conceptual-competence account might have a fairly direct semantic or metaphysical upshot after all.

In what follows, my focus will be knowledge, knowledge ascriptions and the concept of knowledge. I don't claim that any point I make generalizes to other psychological states, ascriptions, concepts, or side-effect effects concerning them.

## 2.2. See, this is ESEE

James Beebe and Wesley Buckwalter (2010) showed that the side-effect effect extends to knowledge ascriptions too. They took Knobe's original case and instead of asking participants whether the CEO intentionally brought about the environmental side effect, they asked whether the CEO knew that the program would bring about the side effect. Beebe and Buckwalter observed a striking asymmetry: participants were significantly more likely to ascribe knowledge if the environment is harmed rather than helped.[2] In other words, they discovered ESEE. Their result has been replicated and appears to be robust (Beebe and Jensen 2012; see also section 4.1 below).

Beebe and Buckwalter conjecture that ESEE might reveal something important about the nature of knowledge and its relation to moral value: whether a belief counts as knowledge can depend on the moral status of actions based on it.

---

[2]  Right around the same time, some armchair epistemologist got lucky and predicted the same thing (see Turri 2012a).

But they are careful to note that there is no easy path from the observed patterns of knowledge attribution among laypeople to sound theoretical conclusions about the nature of knowledge. Beebe and Jensen (2012) are likewise cautious about inferring anything about the nature of knowledge itself from the ESEE data. But Beebe, Buckwalter and Jensen all seem to agree that the ESEE data lend support to a view about the nature of knowledge which has lately been gathering support, namely, that factors other than truth, belief, evidence and reliability can directly affect whether you know (e.g. Fantl and McGrath 2002, 2009, Hawthorne 2004, Stanley 2005). On this approach, practical, moral and other traditionally *nonepistemic* evaluative factors can directly affect whether you know. This is an intriguing thesis which, if true, reveals something deep and important about knowledge.

### 2.3. Gee, this **is** ESEE

Buckwalter (2013) and Beebe and Shea (forthcoming) have recently extended this line of reasoning to "Gettier cases." Very roughly, a Gettier case is a case where a subject has a justified true belief which, according to conventional wisdom in philosophy, fails to count as knowledge due to an improbable confluence of luck (Gettier 1963, Starmans and Friedman 2012, Turri 2011, Turri 2012a, Turri forthcoming b). In particular, Gettier cases involve two strokes of luck: a stroke of bad luck, which would ordinarily prevent the subject's belief from being true, followed by a stroke of good luck, which "cancels out" the bad luck by making the subject's belief true anyway (Zagzebski's 1994: 66; Zagzebksi 1996: 288–9; compare Sosa 1991: 238).

Beebe and Shea asked participants about a CEO very similar to the one in Knobe's original story, except that this CEO has been "Gettiered." Participants were asked to agree or disagree with the claim that the CEO knew that the environment would be helped or harmed. Participants rated their agreement or dis-

agreement on a 7-point Likert scale. They were significantly more likely to agree with the statement if the side effect was harmful rather than helpful. The mean response (4.35) fell above the mid-point and responses of 5 and 7 were both modes, indicating that participants tended to think that the Gettiered CEO knew. Beebe and Jensen observed similar results with different cover stories. Buckwalter has observed even more striking results with some of his vignettes, with mean responses topping 6 on a 7-point scale. This is *Gettier made ESEE*, or the *Gettier epistemic side-effect effect* (GESEE). I replicated these results (see section 4.2 below).

A natural reaction to these cases is that participant response is being distorted by a desire to blame the protagonist and to hold him accountable for a rotten outcome that he could have prevented (compare Alicke 2008, Nadelhoffer 2006). Agreeing that the CEO 'knew' is a way of agreeing that he should be held responsible for the environmental harm. By contrast, people aren't interested in giving the CEO credit for environmental improvement that he didn't even care about. But Buckwalter (2013) has observed the effect persist even when participants are asked whether some *third-party*, who can't plausibly be held responsible, knows that the environmental side effect will occur. So a straightforward blame-based distortion account doesn't explain all the data.

If we're attracted to the idea that our ordinary practice of ascribing knowledge should at least broadly constrain theorizing about knowledge itself, then these data seriously threaten some well entrenched positions in contemporary epistemology. Perhaps whether you know does partly depend on whether you're up to no good. Perhaps we should leave room in our epistemology for Gettiered knowledge.

## 3.  How easy can it get?

Maybe ESEE can be extended to unsettle other articles of conventional wisdom too. Conventional wisdom has it that knowledge requires truth, and that knowledge requires belief. Are these requirements reflected in ordinary usage?

### 3.1. It's not easy being . . . CHESEE?

Consider this contrarian hypothesis:

> *Contrarian Hypothesis* (CH): Knowledge is not factive.

Otherwise put, CH says that knowledge does not require truth: it's possible to know false things. Some philosophers have argued that, on the ordinary conception of knowledge, CH is true (Hazlett 2010, 2012). But virtually all philosophers think that CH is obviously false. And Buckwalter (ms.) has reported several studies which suggest that, on the ordinary conception of knowledge, CH is false. But Buckwalter cautiously reminds us that there might be other evidence that knowledge, ordinarily understood, is not factive. We saw SEE extended to ESEE, which was in turn extended to GESEE. Perhaps the same phenomenon might also provide evidence for CH. In other words, we've seen that knowledge ascriptions can be made ESEE, so can they also be made CHESEE (pronounced 'cheesy')?

### 3.2.  Is knowledge a breeze?

Consider also this unorthodox view:

> *Belief not Required* (BR): Knowledge does not require belief.

Some philosophers have advocated BR (e.g. Radford 1966, Lewis 1996), and experimental philosophers have recently taken up its cause (Schwitzgebel and Myers-Schulz forthcoming; Murray, Sytsma and Livengood 2012; Buckwalter, Rose and Turri, 2013). But most philosophers think that BR is obviously false; they think it's *obvious* that knowledge entails belief (Rose and Schaffer 2013). But per-

haps that's not the ordinary view. Perhaps the ESEE pattern will extend to BR too. We've seen that knowledge ascriptions can be made ESEE. Can they also be made BRESEE (pronounced 'breezy')?

## 4. Actually, this is too easy

Suppose that in the same way that knowledge ascriptions can be made ESEE and GESEE, they can also be made CHESEE or BRESEE or both. What would that show?

I think that it would make a conceptual-competence explanation of the observed effect much less likely. It would seem more likely that participants are incompetently applying their concept of knowledge, or competently but falsely applying it for pragmatic reasons, or competently applying some other concept in response to a 'knowledge' question. Over thousands of years and across many different cultures, careful reflection has led people repeatedly and to the view that knowledge requires truth and belief (or something very similar). This is overwhelmingly reflected in the most influential historical and contemporary theories of knowledge (Matilal 1986, Steup 1996, Lehrer 2000, Feldman 2003, Fumerton 2006, BonJour 2009, Phillips 2011). And if a conceptual-competence explanation is unlikely, then we need to be much more cautious when determining whether the ESEE data support various substantive proposals about the nature of knowledge itself. Of course, none of that would go to show that the effect is unreal. On the contrary, the point is that the effect is, ironically, *too real* for it to guide a theory of knowledge.

By contrast, if knowledge ascriptions can't be made CHESEE or BRESEE in that same way, it eliminates two barriers to a conceptual-competence account. Of course, it would not entail that a conceptual-competence account is true. But it would eliminate potential worries about the effect not ruled out by previous stud-

ies.

## *4.1. Experiment 1*

This section reports an experiment designed to test whether knowledge ascriptions can be made CHESEE. That is, the experiment tests whether the method used to detect ESEE and GESEE in earlier experiments will also produce results that support the unorthodox view that knowledge doesn't require truth.

### *4.1.1. Method*

Participants ($N$ = 147)[3] were randomly assigned to one of six conditions in a 2 (Valence of side-effect: Help/Harm) x 3 (Luck: Normal/False/Gettier) between-subjects experiment: Normal Harm, False Harm, Gettier Harm, Normal Help, False Help, Gettier Help. (See Table 1.) Each participant read one of six stories.

The stories for the conditions were variants of the CEO case, built up from narrative modules. These were the narrative modules (help/harm variations underlined and separated by a slash):

> (*No Luck*) The vice president of a manufacturing company went to the CEO and said, "We are thinking of starting a new program. It will increase our profits, and/but it will also improve/harm local water quality, because it requires that we start/stop dumping polluted water into the river behind our manufacturing plant." The CEO answered, "I don't care at all about improving/harming local water quality. I just want to make as much profit

---

[3] Sixty-seven female, aged 18–65, $M$ = 28.8, $SD$ = 9.92. As with the experiments reported below, participants were recruited using Amazon Mechanical Turk and compensated $0.25 for approximately 2–3 minutes of their time. Participants were not allowed to re-take any survey reported here, and participants who had taken previous similar surveys were excluded by their AMT Worker ID. Participants were located throughout the United States. Ninety-eight percent reported English as their native language. They filled out a brief demographic survey after testing. I excluded data from fifty-five participants who failed comprehension questions. Including data from these participants made a small but statistically insignificant difference to the results reported below. Previous studies of ESEE effects don't report asking comprehension questions.

as I can." Then the CEO made his decision: "Let's start the new program and make some serious profits." The company then started the new program, and sure enough, over the next few months, it caused their profits to increase dramatically.

(*Bad Luck*) But as it turns out, the vice president was wrong about something: the company's new program did not require them to start/stop dumping polluted water into the river. So they never did start/stop.

(*Good Luck – Harm*) However, at the same time that the company was starting its new program, a local military installation decided to save money on expensive waste disposal. Instead of going through the proper procedures, they decided to secretly dump all of their toxic waste straight into a local lake, which harmed local water quality. This illegal and harmful dumping remained a secret and was never made public. When the CEO made his decision to start the new program, he was unaware that the military was going to dump the toxic waste.

(*Good Luck – Help*) However, at the same time that the company was starting its new program, the mayor decided to take action and improve local water quality. Instead of going through the proper procedure, the mayor secretly ordered the local water department to upgrade the filtration system at the local water treatment plant, which improved local water quality. This illegal but helpful expenditure remained a secret and was never made public. When the CEO made his decision to start the new program, he was unaware that the mayor was going to order the upgrade.

 Table 1 shows how stories for the different conditions were built, along with the questions used. The "Good Luck" and "Bad Luck" modules combine to form Gettier cases by virtue of the "double luck" recipe described earlier (section 2.3).

These were the comprehension questions (options in brackets):[4]

(CQ1) Did the company's program cause their profits to increase? [Yes/No]

(CQ2) Did the company's program improve/harm local water quality? [Yes/No]

(CQ3) When the CEO makes his decision to start the program, he thinks that local water quality will be improved/ harmed because he thinks that _____. [the CEO's company will start/stop dumping pollution into the river/the military will start dumping toxic waste into the lake/the local water treatment plant will be upgraded]

(CQ4) What causes local water quality to be improved/ harmed? [The CEO's company started/stopped dumping pollution into the river./The military started dumping toxic waste into the lake./The local water treatment plant was upgraded.][5]

The test question was:

(KLikert) Please tell us the extent to which you agree or disagree with the following statement: At the point in the story when the CEO makes his decision to start the program, he knows that local water quality will be harmed.

Responses were collected on a five-point scale: strongly disagree (= 1), disagree, neutral, agree, strongly agree (= 5). Questions were all presented on a single screen and the story remained at the top of the screen throughout. Ques-

---

[4] Options were rotated randomly for all questions in all experiments reported here, except for Likert scales and confidence measures, which were always ordered low-to-high. Likewise, questions were always presented in the same order in all experiments.

[5] CQ4 was not asked in the False Harm and False Help conditions because it confusingly presupposes that local water quality was improved/harmed. The stories for these conditions give no indication that anything does improve/harm local water quality.

tions were always presented in the same order. Response options for the questions were rotated randomly, except for the options on Likert scales, which were always presented in the same order.[6]

**Table 1**: Experiment 1: List of modules composing the storylines, along with the questions, in the different conditions.

| Normal Harm | False Harm | Gettier Harm |
|---|---|---|
| No Luck<br>CQ1–4<br>KLikert | No Luck<br>Bad Luck<br>CQ1–3<br>KLikert | No Luck<br>Bad Luck<br>Good Luck — Harm<br>CQ1–4<br>KLikert |
| **Normal Help** | **False Help** | **Gettier Help** |
| No Luck<br>CQ1–4<br>KLikert | No Luck<br>Bad Luck<br>CQ1–3<br>KLikert | No Luck<br>Bad Luck<br>Good Luck — Help<br>CQ1–4<br>KLikert |

Setting up the experiment this way allowed me to accomplish three things at once. First, comparing the Normal conditions will test whether we can replicate earlier experiments in which participants were more likely to count a true belief as knowledge in the Harm condition than in the Help condition; that is, it tests ESEE. Second, comparing the Gettier conditions will test whether we can replicate earlier experiments in which participants were more likely to count a Gettiered belief as knowledge in the Harm condition than in the Help condition; that is, it tests GESEE. Third, comparing the False conditions will test whether participants are more likely to count a false belief as knowledge in the Harm condition than in the Help condition; that is, it tests CHESEE.

### 4.1.2. Results and discussion

A one-way analysis of variance (ANOVA) revealed a main effect of Valence[7] and

---

[6]  The same is true in all experiments reported here.

[7]  $F(1, 141) = 34.9$, $p < .001$, $\eta_p^2 = .198$, all tests two-tailed unless otherwise noted.

an interaction effect between Valence and Luck.[8] There was no main effect of Luck.[9]

The results from the Normal conditions replicated earlier findings on ESEE. Knowledge scores in Normal Harm were higher than in Normal Help.[10] The mean knowledge score in both conditions was significantly above the midpoint.[11] The mode in Normal Harm was 5, compared to a mode of 4 in Normal Help.

The results from the Gettier conditions replicated earlier findings on GESEE. Knowledge scores in Gettier Harm were trending higher than in Gettier Help.[12] The mean knowledge score in both conditions was significantly above the mid-point.[13] The mode in Gettier Harm was 5, compared to a mode of 4 in Gettier Help.

The results from the False conditions provide equally good evidence for CHE-SEE. Knowledge scores in False Harm were higher than in False Help.[14] The mean knowledge score in the Harm condition was significantly above the mid-point[15] and virtually identical to Normal Harm. The mode in False Harm was 5, compared to a mode of 4 in False Help.

Figure 1 visualizes the results from this experiment.

---

[8]  $F(2, 141) = 5.2$, $p = .013$, $\eta_p^2 = .059$.

[9]  $F(2, 141) = 1.48$, $p = .231$, n.s.

[10]  Normal Harm, $M = 4.79$, $SD = 0.77$; Normal Help, $M = 3.77$, $SD = 1.24$; $t(40) = 3.28$, MD $= 1.02$, $p = .002$.

[11]  Normal Harm, $t(28) = 12.5$, $p < 0.001$; Normal Help, $t(12) = 2.245$, $p = 0.044$; test value = 3.

[12]  Gettier Harm, $M = 4.24$, $SD = 1.23$; Gettier Help, $M = 3.74$, $SD = 1.15$; $t(54) = 1.56$, MD = $0.5$, $p = .0625$, one-tailed.

[13]  Gettier Harm, $t(24) = 5.02$, $p < .001$; Gettier Help, $t(30) = 3.6$, $p = .001$; test value = 3.

[14]  False Harm, $M = 4.76$, $SD = 0.83$; False Help, $M = 3.0$, $SD = 1.25$; $t(39.76) = 5.824$, MD = $1.76$, $p < .001$.
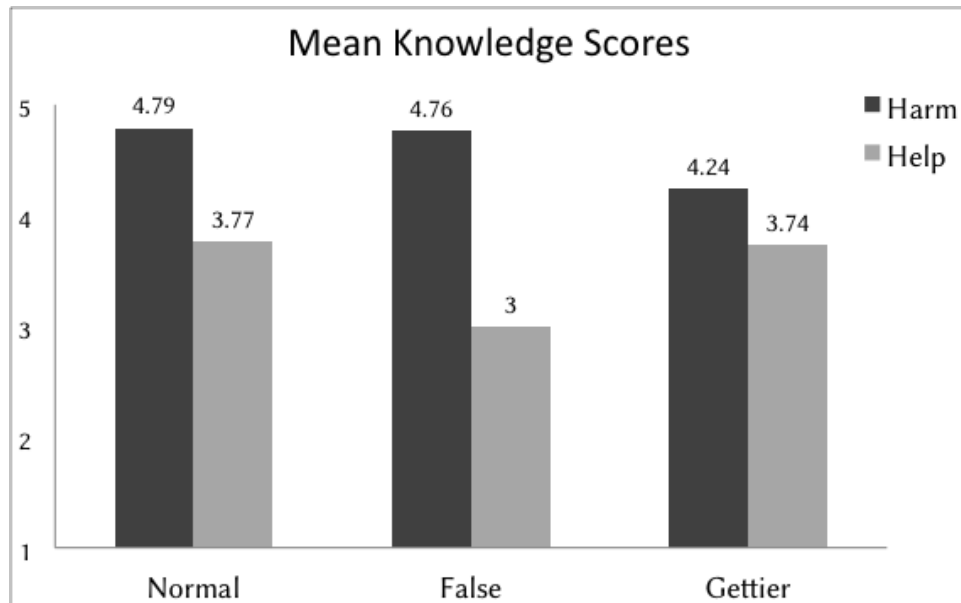
[15]  $t(24) = 10.6$, $p < 0.001$; test value = 3.

**Figure 1**: Experiment 1: Mean agreement with the knowledge ascription, on a 1 (strongly disagree) to 5 (strongly agree) scale.

## *4.2. Experiment 2*

This section reports an experiment designed to test whether knowledge ascriptions can be made BRESEE. That is, the experiment tests whether the method used to detect ESEE, GESEE and CHESEE in earlier experiments will also produce results that support the unorthodox view that knowledge doesn't require belief.

### *4.2.1.* Method

Participants ($N$ = 56)[16] were randomly assigned to one of two conditions: Harm and Help. Participants read a minor variant of a CEO case, in which the CEO starts out by saying, 'I agree that this plan will make us lots of profit. But I reject your claim that it will improve/harm local water quality.' Here is the full vignette (help/harm variations underlined and separated by a slash):

---

[16] Twenty female, aged 18–62, $M$ = 28.5, $SD$ = 10.75. I eliminated data from 14 participants who failed comprehension questions. Ninety-nine percent reported English as a native language.

The vice president of a manufacturing company went to the CEO and said, "We are thinking of starting a new program. It will increase our profits, <u>and</u>/<u>but</u> it will also <u>improve</u>/<u>harm</u> local water quality, because it requires that we <u>start</u>/<u>stop</u> dumping polluted water into the river behind our manufacturing plant." The CEO answered, "I agree that this plan will make us lots of profit. But I reject your claim that it will <u>improve</u>/<u>harm</u> local water quality. Anyway, I don't care at all about <u>improving</u>/ <u>harming</u> local water quality. I just want to make as much profit as I can." Then the CEO made his decision: "Let's start the new program and make some serious profits." ¶[17] The company then started the new program, and sure enough, over the next few months, they made lots of profits. But the vice-president was right: local water quality was also <u>improved</u>/<u>harmed</u>.

The comprehension questions were:

(CQ1) Did the company's program cause water quality to be harmed/improved? [Yes/No]

(CQ2) When the CEO makes his decision to start the new program, he _____ the vice president's claim that it will improve/harm local water quality. [rejects/accepts]

The test question was the same as in Experiment 1.

### 4.2.2. Results and discussion

The CEO in each version of the case rejects the claim that the program will harm local water quality. If the earlier pattern of ESEE results holds, then there will be an effect of condition, with Knowledge scores higher in the Harm condition.

This is exactly what was observed. Knowledge scores were higher in the Harm condition than in the Help condition.[18] The mean knowledge score was sig-

---

[17] Indicates paragraph break as it appeared on the participant's screen.
[18] Harm, $M = 3.6$, $SD = 1.36$; Help, $M = 2$, $SD = 1.05$; $t(50.38) = 4.64$, MD = 1.6, $p < .001$.

nificantly above the midpoint in Harm[19] and significantly below the midpoint in Help.[20] The modes in Harm were 4 and 5, compared to a mode of 2 in Help.

It might be thought that participants in Harm were attributing to the CEO belief in the relevant claim. But this seems unlikely because there was a comprehension question (CQ2) to eliminate such participants. Only participants who answered that the CEO rejects the claim were included in the analysis. Moreover, there was no effect of condition on whether participants failed CQ2.[21] Figure 2 (section 4.4.2 below) visualizes the results from this experiment.

## 4.3. Experiment 3

We've seen that participants agree that the CEO knows when the proposition in question is false, and when the CEO doesn't believe the proposition. But suppose that the proposition is both false and not believed. Will participants then disagree with the knowledge ascription? This section reports an experiment that answers that question. (Unfortunately, I was unable to invent a memorable acronym for this effect.)

### 4.3.1. Method

Participants ($N = 57$)[22] were randomly assigned to one of two conditions: Harm and Help. Participants in this experiment received the exact same treatment as their counterparts in Experiment 2, except that the final sentence of the story was, "But the vice-president was wrong about something: local water quality was not improved/harmed."

---

[19] $t(34) = 2.62, p = .013.$
[20] $t(20) = -4.37, p < .001.$
[21] Fisher's exact test, $p = .517$
[22] Twenty-nine female, aged 18–71, $M = 26.9, SD = 8.84$. One hundred percent reported English as a native language. I excluded data from fifteen participants who failed comprehension questions.

### 4.3.2. Results and discussion

Once again knowledge scores were higher in Harm than in Help.[23] The mean knowledge score didn't differ significantly from midpoint in Harm,[24] whereas it was significantly below the midpoint in Help.[25] The modes in Harm were 2 and 4, compared to a mode of 1 in Help.

## 4.4. Experiment 4

I went even further in the quest to get participants to disagree with the knowledge ascription. I presented them with stories that take away not only truth or belief, but also justification and, in one case, *all three*: belief, truth and justification. I dispensed with the comparison to a Help condition at this point. Would participants finally disagree?

### 4.4.1. Method

Participants ($N$ = 47)[26] were randomly assigned to one of two conditions: True and False. Each participant read one of two stories. In each story, the CEO is not justified in believing that the program will harm local water quality. The vice-president says,

> We are thinking of starting a new program. It will increase our profits, but there is a very, very small chance that it will harm local water quality, since it is just possible that it will require us to dump polluted water into the river behind our manufacturing plant. Nevertheless, that almost certainly won't happen.

The CEO then explicitly rejects the claim that *it is even possible* that the program

---

[23] Harm, $M$ = 2.94, $SD$ = 1.21; Help, $M$ = 1.73, $SD$ = 1.28; $t(55)$ = 3.6, MD = 1.21, $p < .001$.
[24] $t(34)$ = -0.279, $p = .782$, n.s., test value = 3.
[25] $t(21)$ = -4.67, $p < .001$, test value = 3.
[26] Twenty-three female, aged 19–65, $M$ = 29, $SD$ = 10.2. Ninety-five percent reported English as a native language. I eliminated data from fourteen participants who failed comprehension questions.

will harm water quality. Participants in the True condition read a story in which it nevertheless turns out that, 'against all odds,' the program harms water quality. Participants in the False condition read a version in which there 'was never a real chance' that the program would harm water quality. (The full vignette is included in Appendix A.) The comprehension and test questions were the same as in Experiment 2.

### 4.4.2. Results and discussion

A one-way ANOVA revealed no effect of condition on response to the knowledge attribution.[27] In neither condition did mean response differ significantly from the midpoint.[28] The mode in False was 2 (followed closely by 5!), while 2 and 5 were both modes in True.

In the story for the False condition in this experiment, the CEO isn't justified in believing the claim in question, he rejects the claim, and the claim is false. But participants still were *overall neutral* on whether this **unjustified false non-belief** is knowledge! Perhaps if they were told that the CEO had died last week before the proposed program had even been invented, then they would unambiguously disagree that he knows. Figure 2 visualizes the results from Experiments 2–4.

---

[27] $F(1, 45) = 0.714$, $p = .403$, n.s.
[28] True, $M = 3.25$, $SD = 1.54$, $t(23) = 0.796$, $p = .434$, n.s.; False, $M = 2.87$, $SD = 1.55$, $t(22) = -0.13$, $p = .690$, n.s.; test value = 3.

*Figure 2*: Experiments 2–4: Mean agreement with the knowledge ascription, on a 1 (strongly disagree) to 5 (strongly agree) scale. Experiment 2: true non-belief; Experiment 3: false non-belief; Experiment 4: unjustified true/false non-belief. No Help condition was included in Experiments 3 and 4.

## 5. Re-evaluating

Experiments 1–3 show that participants are consistently more willing to agree that the CEO knows that the environmental side effect will occur when it is harmful rather than helpful. This is true whether the CEO has a normal justified true belief, a Gettiered belief, a false belief, no belief, and even when the CEO has no belief and the proposition in question is false. What's more, participant agreement with the knowledge ascription was very high in the case of a false belief. Participants in Harm conditions were neutral on whether false non-beliefs were knowledge, on whether unjustified true non-beliefs were knowledge, and even on whether unjustified false non-beliefs were knowledge.

If we take these results at face value, then there is virtually nothing we can do to make participants *disagree* that the CEO knows that the harmful side effect will occur. This poses a serious challenge to the conceptual-competence account

of epistemic side-effect effects, for reasons noted at the beginning of section 4. It seems unlikely that participants in these studies are competently and literally applying the concept of knowledge. A more likely explanation is that a seriously negative reaction to the CEO is causing performance errors, or that at least many participants are agreeing to a statement other than the one explicitly featured in the test question, perhaps along the lines of, 'The CEO is a world-class jerk.'[29] It also seems unlikely that participants are competently and literally applying the concept of *confidently held belief*, since they persist in agreeing that the CEO knows even when participants explicitly acknowledge that the CEO rejects the claim.

## 6. *The big ESEE: a different approach*

I'm interested in trying a different approach to detecting epistemic side-effect effects. It would be good to detect an effect in such a way that it stands a fighting chance of revealing something about our conceptual competence in ascribing knowledge, and perhaps, in turn, about knowledge itself. This requires not only detecting an effect, but an *appropriately circumscribed* effect.

For this to work, participants must also be willing *deny* knowledge in a certain range of control cases when probed in the same way that produces ESEE. For example, they must be willing to deny that false beliefs are knowledge. More generally, it would help — though it is *not* necessarily required — if we detected ESEE in an overall pattern of knowledge attribution that broadly agrees with mainstream theorizing about knowledge. The agreement doesn't have to be perfect. But if the two are totally at odds, then many epistemologists will suspect that the observed effect is not due to conceptual competence but rather to performance error or more pressing practical concerns, such as blaming the CEO or ex-

---

[29] Compare Alicke and Rose 2010.

pressing disapproval.

This section reports an experiment designed to detect ESEE within such a pattern. The experiment takes a different approach to questioning participants. It features a binary knowledge question whose options are 'really knows' and 'only thinks he knows', along with a confidence measure. This basic approach, pioneered by Christina Starmans and Ori Friedman (2012), has proven effective in the past (see also Turri 2013; but see Cullen 2010 for some cautionary points). It's worth trying here.

## 6.1. Experiment 5

### 6.1.1. Method

Participants ($N$ = 228)[30] were randomly assigned to one of six conditions (Table 2) in the same 2 (Valence: Help/Harm) x 3 (Luck: Normal/False/Gettier) between-subjects design as in Experiment 1. (See Table 2.) The stories for the conditions were again built up from narrative modules. Table 2 shows how the stories were built, along with the questions used. The narrative modules were the same as in Experiment 1, as were the comprehension questions. But the test question was different:

(KQ) When the CEO makes his decision to start the new program, he _____ that local water quality will be im-proved/harmed [really knows/only thinks he knows].

Participants were then asked to rate how confident they were in their answer to the test question, 1 (not at all confident) – 10 (completely confident).

---

[30] Eighty-three female, aged 18–67, $M$ = 27.2, $SD$ = 9.75. Ninety-seven percent reported English as a native language.

**Table 2**: Experiment 5: List of modules composing the storylines, along with the questions.

| Normal Harm | False Harm | Gettier Harm |
|---|---|---|
| No Luck<br>CQ1–4<br>KQ | No Luck<br>Bad Luck<br>CQ1–2, 4<br>KQ | No Luck<br>Bad Luck<br>Good Luck<br>CQ1–4<br>KQ |
| **Normal Help** | **False Help** | **Gettier Help** |
| No Luck<br>CQ1–4<br>KQ | No Luck<br>Bad Luck<br>CQ1–2, 4<br>KQ | No Luck<br>Bad Luck<br>Good Luck<br>CQ1–4<br>KQ |

As before, the Normal stories are very similar to the original CEO cases. The False stories introduce a twist: it turns out that the environmental side effect doesn't occur. The Gettier stories introduce *two* twists: the new program doesn't produce the environmental side effect, but the something else does, which the CEO is unaware of.

### 6.1.2. Results and discussion

Following Starmans and Friedman (2012), I define a *weighted knowledge ascription* as the product of the answer to the dichotomous knowledge question (really knows = 1; only thinks he knows = -1) and the reported confidence (1–10). Scores for this measure thus fall on a twenty-point scale, ranging from -10 (fully confident knowledge denial) to 10 (fully confident knowledge ascription).

A one-way analysis of variance (ANOVA) detected a main effect of Valence,[31] a main effect of Luck,[32] and an interaction of Valence and Luck[33] on weighted knowledge ascription.

Pairwise comparisons with independent-samples t-tests revealed that mean

---

[31] $F(2, 222) = 37.26$, $p < .001$, $\eta_p^2 = .251$.
[32] $F(1, 222) = 10.54$, $p = .001$, $\eta_p^2 = .045$.
[33] $F(2, 222) = 4.90$, $p = .008$, $\eta_p^2 = .042$.

weighted knowledge ascription was higher in Normal Harm than in Normal Help,[34] higher in False Harm than in False Help,[35] but no different between Gettier Harm and Gettier Help.[36] Mean weighted knowledge ascription was significantly above the midpoint in Normal Harm,[37] no different from midpoint in Normal Help,[38] and significantly below the midpoint in both False conditions and both Gettier conditions.[39]

Looking at the dichotomous knowledge question reveals a similar picture. The rate of knowledge attribution was higher in Normal Harm than in Normal Help,[40] and it was higher in False Harm than in False Help,[41] but it didn't differ between Gettier Harm and Gettier Help.[42] Rate of knowledge attribution was above what could be expected by chance in Normal Harm,[43] no different from chance in Normal Help,[44] and significantly below chance in all the other conditions.[45]

Figure 3 visualizes the results from this experiment.

---

[34] Normal Harm, $M$ = 6.34, $SD$ = 6.8; Normal Help, $M$ = 0.03, $SD$ = 3.82; $t(63.17)$ = 3.45, MD = 6.31, $p$ = 0.001.

[35] False Harm, $M$ = -2.68, $SD$ = 8.64; False Help, $M$ = -6.34, $SD$ = 5.81; $t(68.65)$ = 2.18, MD = 3.67, $p$ = .033.

[36] Gettier Harm, $M$ = -6.67, $SD$ = 6.04; Gettier Help, $M$ = -5.88, $SD$ = 6.35; $t(79)$ = -0.57, $p$ = .571, n.s.

[37] $t(37)$ = 5.733, $p$ < .001, test value = 0.

[38] $t(33)$ = 0.02, $p$ = .984, n.s., test value = 0.

[39] One-sample t-tests, all $ps$ ≤ .057, test value = 0.

[40] Harm: 84%, Help: 50%, Fisher's exact test, $p$ = .002.

[41] Harm: 33%, Help: 11%, Fisher's exact test, $p$ = .051.

[42] Harm: 15%, Help: 14%, Fisher's exact test, $p$ = 1.

[43] Binomial test, $p$ < .001.

[44] Binomial test, $p$ = 1.
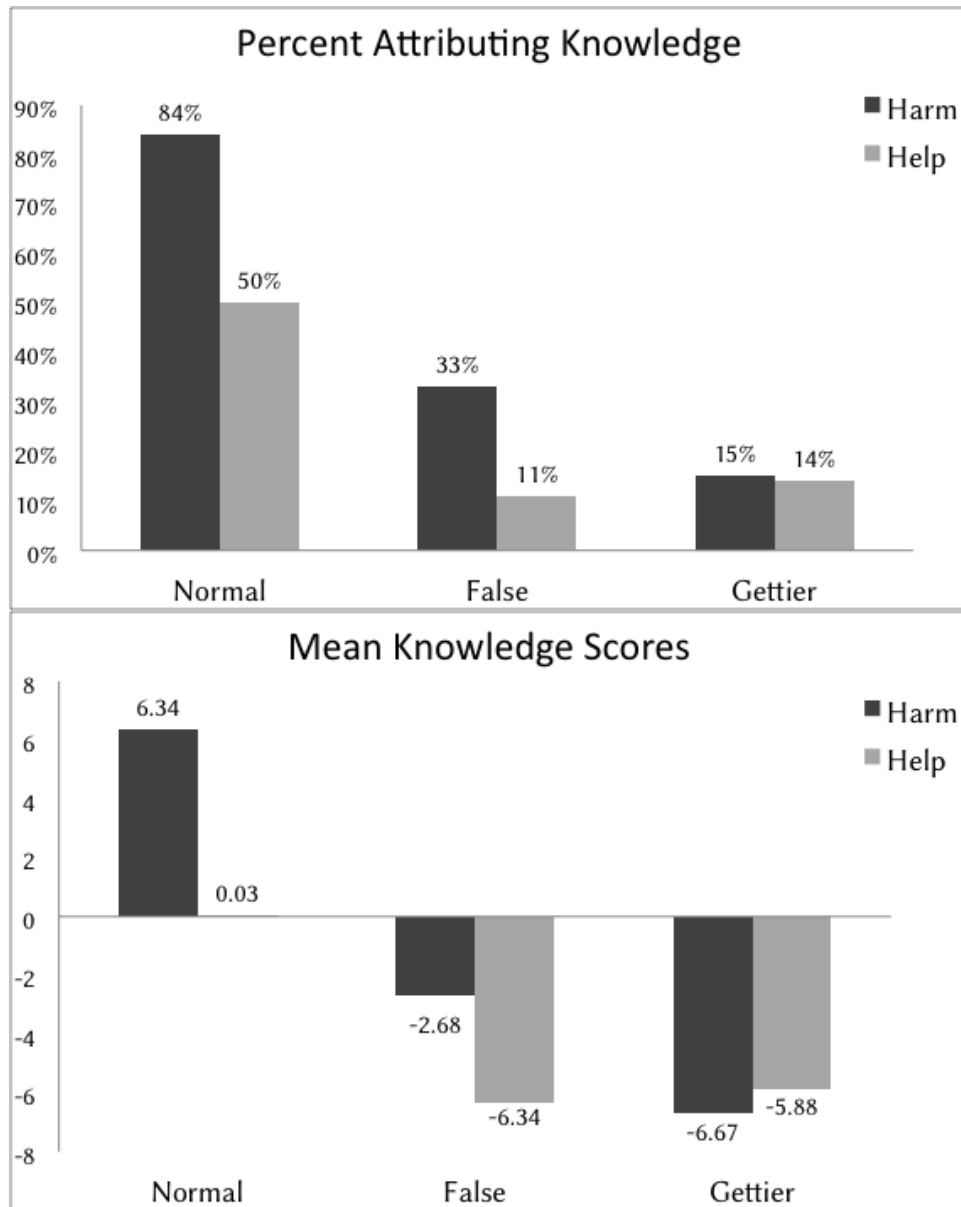
[45] Binomial tests, $ps$ ≤ .038.

**Figure 3**: Experiment 5: *Top panel*: percentage of participants ascribing knowledge in response to the dichotomous question. *Bottom panel*: Mean weighted knowledge ascriptions.

These results provide evidence of an appropriately constrained epistemic side-effect effect. Overall, only one condition produced results that are arguably inconsistent with a mainstream, nonskeptical theory of knowledge: Normal Help. It is surprising that participants in this condition were ambivalent about whether

the CEO really knows that local water quality will be improved. After all, stopping pollution from being dumped into the water is an extremely reliable way of improving water quality. It's possible that random variation resulted in an unusually skeptical lot being assigned to this condition. But even if we suppose that the population mean for weighted knowledge ascription in Normal Help would be significantly above the midpoint at, say, 4, that would still be significantly below what was observed in Normal Harm.[46] So, ultimately, I don't think that the unexpectedly low number in Normal Help should make us especially suspicious of the overall result.

The results from False Harm are also a bit surprising, with 33% saying that the CEO really knows something that is false. Previous studies have shown that roughly 10–15% of participants are typically willing to ascribe false knowledge when the belief is justified (Starmans and Friedman 2012, Buckwalter ms, Turri 2013), but the percentage observed here is significantly greater than even 15%.[47] By contrast, the percentage ascribing false knowledge in False Help is very similar to what has been observed previously in the literature.

Despite these two unexpected observations, there are no intolerable "red flags" in these results that should definitely make us reject the suggestion that participants are competently and literally applying their concept of knowledge. The overall results are mainly consistent with what mainstream theorists of knowledge would predict (again, the one exception being Normal Help). Yet we still observe a significant effect of side-effect valence. A constrained effect of this sort is arguably eligible to inform substantive theorizing about knowledge or our concept thereof.

---

[46] $t(37) = 2.12$, $p = .041$, one-tailed, test value = 4.
[47] Binomial test, $p = .004$, one-tailed, test proportion = .15.

## 7. Conclusion

By this point, it is definitely beyond reasonable doubt that the side-effect effect extends to knowledge ascriptions. ESEE is real and interesting in its own right. But what can it teach us? Ideally it would teach something about our concept of knowledge, and ultimately about knowledge itself. In order for it to teach us either of those things, the effect must manifest itself in a way that is plausibly due to a competent and literal application of our concept of knowledge. And in order to do that, it must be *appropriately constrained*. Up until now, that hasn't appeared to be the case (Experiments 1–4). But Experiment 5 provides evidence of an appropriately constrained effect.

My focus has been whether ESEE is appropriately constrained. But an equally important question for future research is whether it is *appropriately unconstrained*. At least three challenges lie ahead in this direction. I will briefly describe them here.

First, if our concept of knowledge, or knowledge itself, really is sensitive to evaluative facts, then this sensitivity probably won't be limited to side-effect propositions. A *side-effect proposition* is a proposition about a course of action's side effects. For example, in the CEO case above, the side-effect proposition is that local water quality will be harmed or helped. A *central-effect proposition* is a proposition about the primary effects that an action is intended to achieve. In the CEO case, the central-effect proposition is that the new program will increase company profits. It's hard to believe that our conceptual competence in ascribing knowledge would be sensitive to the moral valence of side effects but not to the moral valence of central effects. Similarly, if knowledge itself does depend on non-epistemically evaluative facts in surprising ways, it seems unlikely that this would be restricted to knowledge of side effects. Indeed, knowledge of many sorts of propositions, not necessarily connected to the consequences of the agent's ac-

tions, will probably likewise depend evaluative facts.

Second, if our concept of knowledge, or knowledge itself, really is sensitive to evaluative facts, then this sensitivity probably won't be limited to knowledge based on testimony, or "second-hand knowledge." We should expect to see the effect in cases of "first-hand knowledge" too, such as beliefs based on perception, introspection, inference or memory. Perhaps this second challenge could be resisted on the grounds that different doxastic sources have very different profiles in folk epistemology,[48] in which case other factors might prevent a similar pattern for other sources.

Third, if our concept of knowledge, or knowledge itself, really is sensitive to evaluative facts, then this sensitivity won't be limited to a specific genre of cases. For example, it won't be limited to CEO cases. We should observe an appropriately constrained effect when using a wide range of cover stories. Prior work on ESEE has observed it using multiple cover stories, so I expect that this challenge will be met.

Those are three further tests for understanding ESEE's theoretical significance. If we don't observe an appropriately unconstrained effect, then ESEE might not manifest conceptual competence. By contrast, if we do observe a similar effect for central-effect propositions and first-hand knowledge and a wide range of cover stories, then three things follow. First, the effect will need a new name, since it would have nothing special to do with side effects. (Perhaps *the evaluative effect* will suffice.) Second, and more importantly, a conceptual-competence explanation of the effect becomes much more attractive. Third, consequently, substantive theorizing about the nature of knowledge will arguably have to take the effect into account.

Suppose those challenges are met and we're convinced that ESEE reveals something important about knowledge and our concept thereof. Then we'll face

---

[48]  See Friedman and Turri, under review; Turri, under review a; and Turri, under review b.

an important set of further questions, including *why* knowledge is sensitive to these evaluative considerations, *how* this sensitivity is reflected in the psychology of knowledge attributions, and whether, upon reflection, our concept knowledge *should* be sensitive to such considerations. Why would knowledge work this way? How is that sensitivity reflected in the way people think about and ascribe knowledge? And once we are made aware that our ordinary practice of ascribing knowledge is sensitive in this way, is this something we should endorse, or should we instead change our concept so that it *isn't* thus sensitive, or adopt a related concept that lacks such sensitivity? Some important progress has already been made along these lines by experimental philosophers and naturalists in epistemology (e.g. Schaffer 2008, Knobe 2010, Schaffer and Knobe 2012, Craig 1990, Kornblith 2002, Hawthorne 2004). But this research program would benefit tremendously from greater involvement from other areas of cognitive science besides epistemology and experimental philosophy, including cognitive anthropology, cognitive ethology, and cognitive, evolutionary and developmental psychology.[49]

## References

Adams, Fred and Annie Steadman. 2004. Intentional action in ordinary language: core concept or pragmatic understanding? *Analysis* 64.2: 173–181.

Alexander, Joshua and Jonathan Weinberg. 2007. Analytic epistemology and experimental philosophy. *Philosophy compass* 2.1: 56–80.

Alicke, Mark D. 2008. Blaming badly. *Journal of cognition and culture* 8: 179–186.

Alicke, Mark D. and David Rose. 2010. Culpable control or moral concepts? *Behavioral and brain sciences* 33.4: 330–1.

Austin, J.L. 1956–57. A plea for excuses. *Proceedings of the aristotelian society*, new series, 57: 1–30.

Bach, Kent. 2008. Applying pragmatics to epistemology. *Philosophical issues* 18: 68–88.

Beebe, James and Wesley Buckwalter. 2010. The epistemic side-effect effect. *Mind and language* 25.4: 474–498.

Beebe, James and Mark Jensen. 2012. Surprising connections between knowledge and action: the robustness of the epistemic side-effect effect. *Philosophical psychology* 25.5: 689–715.

Beebe, James and Joseph Shea. Forthcoming. Gettierized Knobe effects. *Episteme*.

BonJour, Lawrence. 2009. *Epistemology: classical problems and contemporary responses*, 2 ed. Rowman and Littlefield.

Buckwalter, Wesley. 2012. Non-traditional factors in judgments about knowledge. *Philosophy compass* 7: 278–289.

Buckwalter, Wesley. 2013. Gettier made ESEE. *Philosophical psychology*.

Buckwalter, Wesley. Unpublished ms. Factive verbs and protagonist projection.

Buckwalter, Wesley, David Rose and John Turri. 2013. Belief through thick and thin. *Noûs*. doi:10.1111/nous.12048

Burge, Tyler. 1979. Individualism and the mental. *Midwest studies in philosophy* 4: 73–121.

Burge, Tyler. 1986. Individualism and psychology. *Philosophical review* 95: 3–45.

Cohen, Stewart. 1999. Contextualism, skepticism, and the structure of reasons. *Philosophical perspectives* 13: 57–89.

Craig, Edward. 1990. *Knowledge and the state of nature: an essay in conceptual synthesis*. Oxford UP.

Cullen, Simon. 2010. Survey-driven romanticism. *Review of philosophy and psychology* 1: 275–296.

Davidson, Donald. 1983. A coherence theory of truth and knowledge. Reprinted in *Subjective, intersubjective, objective*. Oxford UP, 2001.

Davis, Wayne. 2007. Knowledge claims and context: loose use. *Philosophical studies* 132: 395–438.

DeRose, Keith. 2009. *The case for contextualism: knowledge, skepticism, and context, volume 1*. Oxford UP.

Fantl, Jeremy and Matthew McGrath. 2009. *Knowledge in an uncertain world*. Ox-

ford UP.

Feldman, Richard. 2003. *Epistemology*. Prentice Hall.

Feltz, Adam and Chris Zarpentine. 2010. Do you know more when it matters less? *Philosophical psychology* 23: 683–706.

Friedman, Ori and John Turri. Under review. Is probabilistic evidence a source of knowledge?

Fumerton, Richard. 2006. *Epistemology*. Blackwell.

Gettier, Edmund. 1963. Is justified true belief knowledge? *Analysis* 23.6: 121–3.

Guglielmo, Steve and Bertram Malle. 2010. Can unintended side effects be intentional? Resolving a controversy over intentionality and morality. *Personality and social psychology bulletin* 36.12: 1635–1647.

Hawthorne, John. 2004. *Knowledge and lotteries*. Oxford UP.

Hazlett, Allan. 2010. The myth of factive verbs. *Philosophy and phenomenological research* 80: 497–522.

Hazlett, Allan. Forthcoming. Factive presupposition and the truth condition on knowledge. *Acta analytica*.

Holton, Richard. 1997. Some telling examples: a reply to Tsohatzidis. *Journal of pragmatics* 28: 624–628.

Lewis, David. 1996. Elusive knowledge. *Australasian Journal of Philosophy* 74: 549–567.

Knobe, Joshua. 2003a. Intentional action and side-effects in ordinary language. *Analysis* 63: 190–193.

Knobe, Joshua. 2003b. Intentional action in folk psychology: an experimental investigation. *Philosophical psychology* 16: 309–24.

Knobe, Joshua. 2004. Intention, intentional action and moral considerations. *Analysis* 64: 181–187.

Knobe, Joshua. 2010. Person as scientist, person as moralist. *Behavioral and brain sciences* 33: 315–329.

Knobe, Joshua and Arudra Burra. 2006. Experimental philosophy and folk concepts: methodological considerations. *Journal of cognition and culture* 6.1–2: 331–342.

Kornblith, Hilary. 2002. *Knowledge and its place in nature*. Oxford UP.

Kripke, Saul. 1972. *Naming and necessity*. Harvard UP.

Lehrer, Keith. 2000. *Theory of knowledge*, 2ed. Westview.

Malle, Bertram. 2006. Intentionality, morality, and their relationship in human judgment. *Journal of cognition and culture* 6.1–2: 87–112.

Matilal, Bimal. 1986. *Perception: an essay on classical Indian theories of knowledge*. Oxford UP.

Murray, Dylan, Justin Sytsma, and Jonathan Livengood. 2012. God knows (but does God believe?). *Philosophical studies*. DOI: 10.1007/s11098-012-0022-5.

Nadelhoffer, Thomas. 2006. Bad acts, blameworthy agents, and intentional actions: some problems for jury impartiality. *Philosophical explorations* 9: 203–220.

Nagel, Jennifer. 2008. Knowledge ascriptions and the psychological consequences of changing stakes. *Australasian journal of philosophy* 86: 279–294.

Pettit, Dean and Joshua Knobe. 2009. The pervasive impact of moral judgment. *Mind & language* 24: 586–604.

Philips, Stephen. 2011. Epistemology and classical Indian philosophy. *Stanford encyclopedia of* philosophy, Spring 2011 edition. Ed. Edward Zalta. Accessed 15 September 2012, <http://plato.stanford.edu/archives/spr2011/entries/epistemology-india/>

Pinillos, Ángel. Forthcoming. Knowledge, experiments and practical interests. *New essays on knowledge ascriptions*. Ed. Jessica Brown and Mikkel Gerken. Oxford UP.

Putnam, Hilary. 1975. The meaning of 'meaning'. *Philosophical papers, v. 2: mind, language, and reality*. Cambridge UP.

Rose, David and Jonathan Schaffer. 2013. Knowledge entails dispositional belief. *Philosophical Studies*.

Rysiew, Patrick. 2001. The context-sensitivity of knowledge attributions. *Noûs* 35.4: 477–514.

Santayana, George. 1923. *Skepticism and animal faith*. Dover, 1955.

Sartwell, Crispin. 1991. Knowledge is merely true belief. *American philosophical quarterly* 28.2: 157–165.

Schaffer, Jonathan. 2008. Knowledge in the image of assertion. *Philosophical issues* 18: 1–19.

Schaffer, Jonathan and Joshua Knobe. 2012. Contrastive knowledge surveyed. *Noûs*. 46: 675–708.

Schwitzgebel, Eric and Blake Myers-Schulz. Forthcoming. Knowing that P without believing that P. *Noûs*.

Sosa, Ernest. 1991. *Knowledge in Perspective: Selected Essays in Epistemology*. Cambridge: Cambridge University Press.

Sripada, Chandra and Jason Stanley. Forthcoming. Empirical tests of interest-relative invariantism. *Episteme*.

Stanley, Jason. 2005. *Knowledge and practical interests*. Oxford UP.

Starmans, Christina and Ori Friedman. 2012. The folk conception of knowledge. *Cognition* 124: 272–283.

Steup, Matthias. 1996. *An introduction to contemporary epistemology*. Prentice Hall.

Swain, Stacey, Joshua Alexander and Jonathan Weinberg. 2008. The instability of philosophical intuitions: running hot and cold on Truetemp. *Philosophy and phenomenological research* 76.1: 138–155.

Turri, John. 2011. Manifest Failure: the Gettier problem solved. *Philosophers' imprint* 11.8: 1–11

Turri, John. 2012a. Is knowledge justified true belief? *Synthese* 184.3: 247–259.

Turri, John. 2012b. In Gettier's wake. *Epistemology: the key thinkers*. Ed. Stephen Hetherington. Continuum.

Turri, John. 2013. A conspicuous art: putting Gettier to the test. *Philosophers' imprint*.

Turri, John. Forthcoming. Linguistic intuitions in context: a defense of pure non-skeptical invariantism. *Intuitions*. Ed. Anthony Booth and Darrell Rowbottom. Oxford UP.

Turri, John. Under review a. An open-and-shut case: epistemic closure in the manifest image.

Turri, John. Under review b. Skeptical appeal: the source-content bias and its application to skepticism.

Weinberg, Jonathan, Shaun Nichols and Stephen Stich. 2001. Normativity and epistemic intuitions. *Philosophical topics* 29: 429–60.

Young, Liane, Fiery Cushman, Ralph Adolphs, Daniel Tranel and Marc Hauser. 2006. Does emotion mediate the effect of an action's moral status on its intentional status? Neuropsychological evidence. *Journal of cognition and*

*culture* 6: 265–278.

Zagzebski, Linda. 1994. "The Inescapability of Gettier Problems." *The Philosophical Quarterly* 44.174: 65–73.

Zagzebski, Linda. 1996. *Virtues of the Mind: An Inquiry into the Nature of Virtue and the Ethical Foundations of Knowledge*. Cambridge: Cambridge University Press.

## *Appendix A: materials for Experiment 4*

The vignette:

> The vice president of a manufacturing company went to the CEO and said, "We are thinking of starting a new program. It will increase our profits, *but there is a very, very small chance that it will harm local water quality, since it is just possible* that it will require us to dump polluted water into the river behind our manufacturing plant. Nevertheless, *that almost certainly won't happen*." The CEO answered, "I agree that this plan will make us lots of profit. But *I reject your claim that it is even possible* that it will harm local water quality. Anyway, I don't care at all about harming local water quality. I just want to make as much profit as I can." Then the CEO made his decision: "Let's start the new program and make some serious profits." ¶ The company then started the new program, and sure enough, over the next few months, they made lots of profits. But the vice-president was wrong about something: there was never a real chance that it would harm local water quality [But, against all odds, the program did harm local water quality].