Gettierized Knobe Effects*

James R. Beebe and Joseph Shea (University at Buffalo)

Episteme (forthcoming)

We report experimental results showing that participants are more likely to attribute knowledge in familiar Gettier cases when the would-be knowers are performing actions that are negative in some way (e.g., harmful, blameworthy, norm-violating) than when they are performing positive or neutral actions. Our experiments bring together important elements from the Gettier case literature in epistemology and the Knobe effect literature in experimental philosophy and reveal new insights into folk patterns of knowledge attribution.

keywords: Knobe effect, Gettier problem, experimental philosophy, experimental epistemology

1. Introduction

David Lewis (1983, p. x) famously remarked, "Philosophical theories are never refuted conclusively. (Or hardly ever, Gödel and Gettier may have done it.)" Edmund Gettier's (1963) landmark contribution to philosophy was responsible for setting the agenda for much of what we now know as contemporary epistemology—e.g., the existing debate over necessary and sufficient conditions for knowledge, epistemic internalism and externalism, relevant alternatives, epistemic luck, and much else besides. In recent years, however, experimental philosophers have begun raising questions about how widely shared the 'Gettier intuition'—viz., the intuition that

* Thanks to Mark Alfano, Brian Robinson, David Sackris, Kevin McCain, an anonymous reviewer for *Episteme*, and audience members at the 2012 meeting of the Central States Philosophical Association for helpful comments on previous versions of this paper.

subjects in Gettier cases lack knowledge—really is.[1]

In the founding document of experimental epistemology, for example, Jonathan Weinberg, Shaun Nichols, and Stephen Stich (2001) presented participants with the following version of one of Gettier's original cases:

> Bob has a friend, Jill, who has driven a Buick for many years. Bob therefore thinks that Jill drives an American car. He is not aware, however, that her Buick has recently been stolen, and he is also not aware that Jill has replaced it with a Pontiac, which is a different kind of American car. Does Bob really know that Jill drives an American car, or does he only believe it?

In response to the question 'Does Bob really know that Jill drives an American car, or does he only believe it?' most of the American college students of European ancestry surveyed gave the more "correct" (or at least typical) philosophical response of 'only believes.' However, many students of East Asian (i.e., Korean, Japanese and Chinese) and South Asian (i.e., Indian, Pakistani, Bangladeshi) descent did not (cf. Table 1). Approximately half of East Asians and more than half of South Asians gave the 'really knows' response.

	Really Knows	**Only Believes**
Westerners	26%	74%
East Asians	53%	47%
South Asians	61%	39%

Table 1. Participant responses to Weinberg, Nichols, and Stich's (2001) Gettier case, divided according to ethnic group.

When epistemic intuitions diverge about concrete cases where it had been previously assumed there would be nearly universal agreement, a challenge is posed to the evidential and

[1] For a recent overview of how the Gettier intuition has been viewed, cf. Turri (2011).

argumentative force of those cases. Weinberg, Nichols and Stich (2001) suggest that "a sizeable group of epistemological projects—a group which includes much of what has been done in epistemology in the analytic tradition—would be seriously undermined if one or more of a cluster of empirical hypotheses about epistemic intuitions turns out to be true." One such hypothesis is that epistemic intuitions vary from culture to culture.

Of course, not everyone has been convinced by these recent results. Some point to the fact that Weinberg, Nichols, and Stich (2001) recruited only 23 East Asians and 23 South Asians for their study, all of which were Rutgers undergraduates. Others note that their results have not been replicated in more than ten years and that no other studies have reported east-west differences in epistemic intuitions. Jennifer Nagel (forthcoming a) observes that the responses of South Asians and East Asians "lie closer to the 50-50 split that one sees when subjects are not interested in a problem and are just answering randomly." Even the authors themselves (personal communication) now have doubts—particularly about whether East Asians understood that the intended meaning of 'American car' was 'American-made car' rather than 'car owned by an American.'

Furthermore, even if Asian participants apprehended the intended meaning of 'American car,' crosscultural differences in participants' familiarity with what falls under this concept can affect the ease with which they make higher-order classification judgments. Twelve year old Alex Tenenbaum (son of philosophers Jennifer Nagel and Sergio Tenenbaum) offers the following case as an excellent illustration of this point:

> Ash has a friend, Brock, who has owned a Torchic for many years. Ash therefore thinks that Brock owns a fire type Pokémon. He is not aware, however, that his Torchic has recently been stolen, and he is also not aware that Brock has replaced it with a Ponyta,

> which is a different kind of fire type Pokémon. Does Ash really know that Brock owns a fire type Pokémon, or does he only believe it?

Most readers unfamiliar with the world of Pokémon who read this example for the first time probably do not find the answer to Alex's question to be intuitively obvious. Most importantly, it should be clear that crosscultural differences in the familiarity participants have with concepts such as 'American car' or 'fire type Pokémon' do not amount to interesting differences in epistemic intuitions.

Simon Cullen (2010) presented Western subjects with the Gettier case reprinted above but instructed them to choose between saying either that Bob knows that Jill drives an American car or that Bob does not know that Jill drives an American car. Cullen correctly notes that 'really knows' seems to express a distinct concept from 'knows' and is perhaps more akin to 'knows with certainty.' When Western participants were offered the dichotomous choice between 'knows' and 'does not know,' 42% chose 'knows'—significantly higher than the percentage of those who chose 'really knows' in the Weinberg, Nichols and Stich (2001) study. Cullen's study shows that conclusions drawn about participants' concept of 'knowledge' should not be drawn from participant responses to questions about 'really knowing,' but it should be kept in mind that this does not undermine conclusions one might want to draw about participants' concept of 'really knowing.'

More recently, the world of experimental philosophy began to buzz with excitement when word spread that Christina Starmans and Ori Friedman (2009) had found a significant gender difference in responses given to the following two Gettier cases:

> Sue is about to do the dishes. She removes her wedding ring and lays it on the counter, alongside a dirty fork. She notices she is out of dish soap, so she locks her apartment and

goes to the store downstairs to buy some. Sue's neighbor Ernest is a bit crazy, and has been spying on Sue through a peephole. While Sue is gone, he picks the lock to her apartment, and takes her wedding ring, replacing it with a cheap plastic ring from a gumball machine. He locks her apartment door, and returns home. Sue has only been gone for 5 minutes, and is now on her way back.

Peter is in his locked apartment, and is reading. He decides to have a shower. He puts his book down on the coffee table. Then he takes off his watch, and also puts it on the coffee table. Then he goes into the bathroom. As Peter's shower begins, a burglar silently breaks into Peter's apartment. The burglar takes Peter's watch, puts a cheap plastic watch in its place, and then leaves. Peter has only been in the shower for two minutes, and he did not hear anything.

Like Weinberg, Nichols, and Stich (2001), Starmans and Friedman asked participants whether each protagonist really knows or only believes that the object in question is on the counter or on the coffee table. Female participants were found to be significantly more likely than males to say that the protagonists really know each of these things. However, the initial excitement over these results was quickly squelched when Starmans and Friedman reported that they have been unable to replicate these results and now consider them to be a fluke.[2] Moreover, Nagel, San Juan, and Mar (forthcoming) recently undertook an investigation of Gettier case intuitions and found no gender-based differences.

Despite the mixed track record of recent experimental investigations into Gettier case intuitions, there are nevertheless reasons to expect that interesting findings can be obtained in this area. Building upon Joshua Knobe's (2003a, 2003b, 2004) groundbreaking work on the

[2] It is troublesome that some experimental philosophers who know that Starmans and Friedman have rejected these results continue to tout them as evidence for demographic differences in epistemic intuitions.

‘Knobe effect,’ Beebe and Buckwalter (2010) and Beebe and Jensen (2012) have found that participants are more likely to think an agent knows a given side-effect will result from their primary action when that side-effect is bad than when it is good—even when the evidence available to the agent in the bad condition is seemingly identical to the evidence available in the good condition. In other words, the goodness or badness of an action or the praiseworthiness or blameworthiness of the agent performing the action appears able to influence epistemic assessments of the agent’s beliefs, even when traditional epistemic factors are held constant. In light of these results, it seems reasonable to expect that Gettier case intuitions can be made to vary by manipulating features of the actions that protagonists are performing in Gettier-style scenarios.[3] In particular, we predict that participants will be more likely to attribute knowledge that an outcome will occur when that outcome is bad or if the agent is blameworthy, even if the agent’s belief about that outcome is ‘Gettierized.’ How large an effect can be generated by such action manipulations remains to be seen.

We also predict that participants will be more likely to attribute knowledge in basic Gettier scenarios than they will be to deny or remain neutral about whether knowledge is present. This prediction is based upon classroom experiences in which students in epistemology courses do not seem to display the Gettier intuition as readily as the overwhelming consensus among epistemologists regarding Gettier cases might lead some philosophers to expect.

In Sections 2 and 3 we report the results of two experiments that test the foregoing predictions. Experiment 1 takes well known Knobe effect cases and Gettierizes them (i.e., manipulates the agents’ epistemic situations in familiar ways), while Experiment 2 takes well

[3] Buckwalter (forthcoming) independently came to the same conclusion and has been experimentally investigating the same set of phenomena. Cf. Turri (2012) for additional discussion of how considerations from the Gettier problem and Knobe effect literatures can be combined.

known Gettier cases and 'Knobifies'[4] them (i.e., manipulates the valence of the believers' actions). The results of both experiments confirm our central hypotheses, revealing further facts about the extent to which folk epistemic intuitions can be influenced by what seem to be epistemically irrelevant factors.

For present purposes, we treat a Gettier case as one in which a subject has a belief that is both justified and true but where elements of epistemic luck have conspired to prevent the justification and truth of the belief from being related in the expected fashion. We do not take a stand on precisely what (if anything) distinguishes a Gettier case from a barn-façade case or from a case of double epistemic luck that is neither a Gettier case nor a barn-façade case. In fact, we suspect there is probably not a principled distinction to be made along these lines. Furthermore, in treating something as a Gettier case we are not insisting that the believer in question fails to have knowledge. Thus, when constructing our own cases or drawing upon extant cases from the literature, we simply follow the general trend in epistemology and call them all Gettier cases, regardless of whether the believers actually fail to have knowledge. We suspect that believers in cases of double epistemic luck of the sort described above do not have knowledge, but our present concerns are officially independent of this issue.

2. Experiment 1

The first case we developed was based upon Knobe's (2003a, p. 191) famous chairman and the environment case:

> ENVIRONMENT: The vice-president of a company went to the chairman of the board and said, "We are thinking of starting a new program. It will help us increase profits, *and/but* it will also *help/harm* the environment." The chairman of the board answered, "I

[4] Thanks to MA for suggesting this term.

> don't care at all about *helping/harming* the environment. I just want to make as much profit as I can. Let's start the new program." They started the new program. Sure enough, the environment was *helped/harmed*.

The first member of each italicized word pair represents the term used in one condition, while each second member represents the term used in the contrasting condition. Knobe presented participants with either the help or the harm version of ENVIRONMENT and asked whether the chairman intentionally helped or harmed the environment. 82% of participants in the harm condition answered that the chairman intentionally harmed the environment, but only 23% of those in the help condition thought he intentionally helped it. We transformed ENVIRONMENT into the following Gettier case by adding elements of epistemic luck that make the would-be knower's belief true for reasons other than what the believer's evidence would ordinarily suggest:

> WATER: The vice-president of a manufacturing company went to the CEO and said, "We are thinking of starting a new program. It will help us increase profits, *and/but* it will also *improve/harm* local water quality." The CEO answered, "I don't care at all about *improving/harming* local water quality. I just want to make as much profit as I can. Let's start the new program." They began implementing the new program, but the vice-president's prediction about *helping/harming* the local water supply turned out to be incorrect. However, shortly after the new program was started, *the city built a new water treatment plant in order to improve water quality*/*another company began to dump toxic waste into the local water supply*. The CEO was unaware of *the city's new plan*/*the actions of this company*. He formed the belief that local water quality levels were going to *rise/fall*. In the coming months, local water quality *rose/fell* significantly.

As before, the first member of each italicized word or phrase pair represents the wording of the first condition, while the second member of each pair represents the wording of the second. As in any good Gettier case, the CEO's belief is both justified—he was given seemingly reliable testimony by the vice-president—and true—the water quality levels did in fact change. However, the change was due to an unforeseen factor. Participants were asked to indicate the extent to which they agreed or disagreed with the following claim: "The CEO knew that local water quality levels were going to *rise/fall*." Answers were reported on a seven-point Likert scale, with '1' marked 'Strongly Disagree,' '4' marked 'Neutral' and '7' 'Strongly Agree.'

Following Knobe (2004) and Beebe and Jensen (2012), a second Knobe effect case that concerned aesthetic (rather than moral) benefits or harms was turned into the following Gettier case pair:

> MOVIES: The Vice-President of a movie studio was talking with the CEO. The Vice-President said: "We are thinking of implementing a new policy. If we implement the policy, it will increase profits for our corporation, *and/but* it will also make our movies *better/worse* from an artistic standpoint." The CEO said: "Look, I don't care one bit about making our movies *better/worse* from an artistic standpoint. All I care about is making as much profit as I can. Let's implement the new policy." They began implementing the new policy, but the Vice-President's prediction about the policy making their movies *better/worse* turned out to be incorrect. However, shortly after the new policy was implemented, another executive at the studio fired all of the movie studio's *inexperienced screenplay writers and replaced them with award-winning writers*/*best screenplay writers and replaced them with inexperienced ones*. The CEO was unaware of this executive's actions. He formed the belief that the artistic quality of

his studio's movies would *improve/go down* soon. In the coming months, the quality of their movies *improved/dropped* significantly.

Again, the CEO's belief is both justified and true, but elements of luck prevent the justification and truth of his belief from being related in the expected fashion. Participants were asked to indicate the extent to which they agreed or disagreed with statement "The CEO knew that the artistic quality of their movies was going to *improve/drop*," and their answers were recorded on the same kind of seven-point scale described above.

Following Knobe and Mendlow (2004) and Beebe and Jensen (2012), a nonmoral Knobe effect case drawn from the business world was Gettierized in the following way:

SALES: Susan is the president of a major computer corporation. One day, her assistant came to her and said, "We are thinking of implementing a new corporate restructuring plan. It will simplify our corporate structure, *and/but* it will also *increase/decrease* sales in New Jersey for the next quarter." Susan replied, "I don't care about what happens in the next quarter. We need to simplify our corporate structure. Let's implement the new plan." They began implementing the new plan, but the assistant's prediction about the plan *increasing/decreasing* sales in New Jersey turned out to be incorrect. However, shortly after the new plan was implemented, Susan's largest clients in New Jersey *decided to upgrade their computers and placed large orders with her corporation/experienced a round of heavy layoffs and budget cutbacks*. Susan was not yet aware of *her clients' decisions/the layoffs and budget cuts*. She formed the belief that sales in New Jersey would *increase/decrease* in the following quarter. In the next quarter, sales in New Jersey *increased/decreased* significantly.

Participants were asked whether they agreed that Susan knew that sales in New Jersey in the next

quarter were going to increase or decrease. Finally, borrowing from Knobe (2007) and Beebe and Jensen (2012), the following pair of Gettierized Knobe effect cases was constructed as well:

> NAZI: In Nazi Germany, there was a law called the "racial identification law." The purpose of the law was to help identify people of certain races so that they could be rounded up and sent to concentration camps. Shortly after this law was passed, the CEO of a small corporation decided to make certain organizational changes. The vice-president of the corporation said: "By making those changes, you'll definitely be increasing our profits. But you'll also be *fulfilling/violating* the requirements of the racial identification law." The CEO said: "I don't care one bit about that. All I care about is making as much profit as I can. Let's make those organizational changes!" As soon as the CEO gave this order, the corporation began making the organizational changes. The vice-president's prediction about *fulfilling/violating* the requirements of the racial identification law turned out to be incorrect. However, shortly after the organizational changes were made, the requirements of the racial identification law were changed, so that the corporation's organizational changes now *fulfilled/violated* those requirements. The CEO was unaware of the recent changes in the law. He formed the belief that his corporation's organizational changes would *fulfill/violate* the law. The changes did in fact *fulfill/violate* the law.

Participants were asked whether they agreed that the CEO knew that the organizational changes would fulfill or violate the requirements of the law.

In a between subjects design 376 undergraduate college students (mean age = 21, 52% female, 64% Anglo-American) from a large, public university in the northeastern United States were each given one of the cases from the four vignette pairs above. Mean participant responses

are represented in Figure 1, with other supporting details in Table 2. There are two reasons the 'rise,' 'improve,' 'increase' and 'fulfill' conditions of the four vignette pairs are all labeled in Figure 1 as varieties of 'fulfill' conditions and each of the contrasting conditions are labeled 'violate' conditions. The first is simply for ease of reference. The second is that each of the actions in the 'fulfill' conditions fulfills some salient norm, while the actions in the 'violate' conditions violate salient norms. Although we believe that the contrast between fulfilling and violating a norm is likely to figure in the ultimate explanation of why participants responded as they did[5], we do not wish to claim at this point that it is the primary explanatory factor.

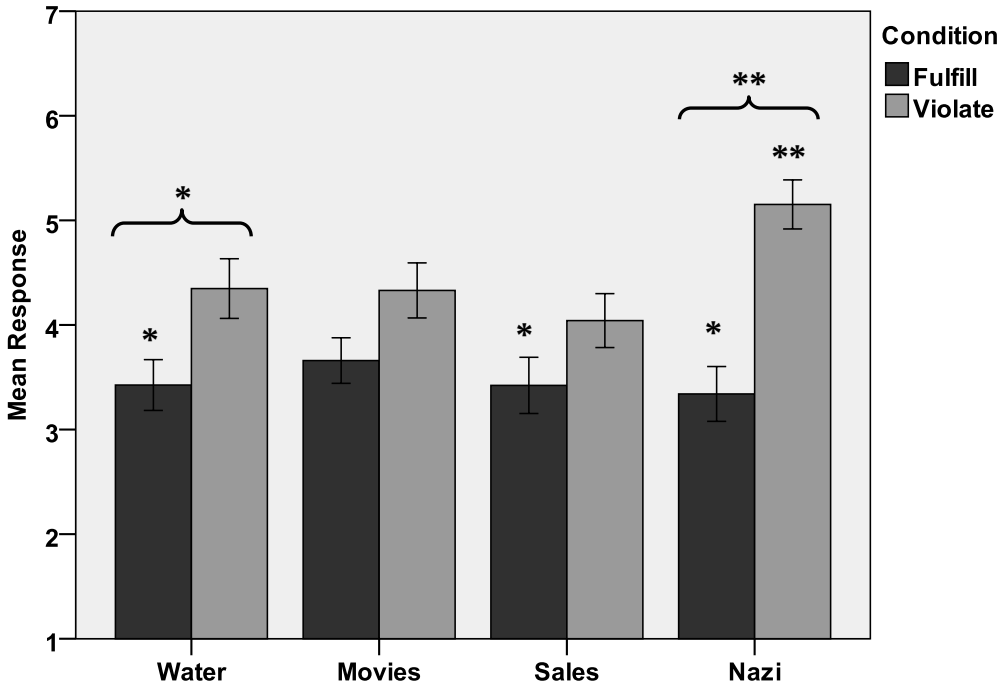

Figure 1. Mean participant responses in each of the eight conditions of Experiment 1. In each graph error bars represent standard errors of the means. An '*' or '**' by itself indicates that the mean differs significantly from the neutral midpoint at either the .05 or the .01 level. An '*' or '**' with a bracket indicates a statistically significant difference between pairs of conditions at either the .05 or the .01 level.[6]

[5] Cf. Alfano, Beebe and Robinson (2012) for details on this kind of approach.

[6] Rise: $t(46) = -2.366$, $p < .05$; Fall: $t(45) = 1.219$, $p > .05$; Improve: $t(45) = -1.317$, $p > .05$; Drop: $t(49) = 1.253$, $p > .05$; Increase: $t(45) = -2.352$, $p < .05$; Decrease: $t(47) = .162$, $p > .05$; Fulfill: $t(46) = -2.516$, $p < .05$;

	WATER		MOVIES		SALES		NAZI	
	Rise	**Fall**	**Improve**	**Drop**	**Increase**	**Decrease**	**Fulfill**	**Violate**
Mean	3.43	4.35	3.72	4.33	3.37	4.04	3.34	5.15
Median	4	4.5	4	5	3	4	3	5
Mode	4	5 & 7	4	5	1 & 2	5	4	5
SE	.24	.29	.22	.26	.27	.26	.26	.24

Table 2. Means, medians, modes, and standard errors for participant responses to each of the four pairs of cases used in Experiment 1.

The first thing to note about the mean responses in Figure 1 and Table 2 is that the first seven out of eight of them lie within .66 of the neutral midpoint of ‘4,’ even though each vignette represents a Gettier case. Table 2 shows that the median response in six of the eight cases is ‘4’ or higher. Granted, the means are not close to ceiling—i.e., participants are not strongly convinced that the protagonists in the Gettier cases have knowledge. But the means, medians and modes are all considerably higher than what traditional philosophical wisdom says is the proper response. Despite the mild complexity of the vignettes, it is clear that what makes the protagonist’s belief true in each case is not what the protagonist is expecting or what the protagonist’s evidence suggests will make it true. As such, according to traditional epistemological thinking, each case should be a case in which the protagonist fails to have knowledge.

Furthermore, from the perspective of almost every epistemologist, participants in each of the four ‘violate’ conditions should not have been more likely to think the central protagonists knew the side-effects in question were going to occur than participants in the contrasting conditions. The most common answer given in the ‘violate’ conditions of MOVIES, SALES and NAZI was ‘5,’ while ‘5’ and ‘7’ were tied for being the most common answer in the ‘violate’

Violate: $t(45) = 4.911$, $p < .001$.

condition of WATER. In each of the four 'violate' conditions more participants selected answers above the neutral midpoint (50%, 52%, 44%, and 74%, respectively) than below it (35%, 28%, 40%, and 15%). However, only in WATER and NAZI was the difference between the means in the 'fulfill' and 'violate' conditions statistically significant, where the effect sizes were small and medium, respectively.[7] Nevertheless, treating Experiment 1 as having a 4 x 2 design and analyzing the results with a two-way ANOVA reveals a main effect for the 'fulfill/violate' variable with a small to medium effect size.[8]

In the 'fulfill' conditions of WATER, MOVIES and NAZI, the most common participant response was '4.' Participants in these conditions chose either to remain neutral (by selecting '4') or to attribute knowledge (by selecting '5,' '6' or '7') 51%, 59%, and 50% of the time. Even in the increase condition of SALES, where '1' and '2' were the most common answers, 46% of participants selected '4,' '5,' '6' or '7.' In other words, in the absence of any 'Knobe effect factor' that might drive up knowledge attributions, participants were not significantly more inclined to deny knowledge than they were to fail to deny it. Some scholars (e.g., Nadelhoffer 2004; Alicke 2008) have argued that participants' comparatively stronger inclinations to attribute certain folk psychological states in Knobe effect cases can be explained to a large extent by (often affect-driven) cognitive processes that are responsible for blame attribution having a distorting effect upon the processes responsible for other mental state attributions. These kinds of considerations, however, cannot be marshaled to explain why participants were as willing to attribute knowledge in the 'fulfill' conditions of Experiment 1 as they were, since the actions performed in these cases are not in general blameworthy.

---

[7] Water: $t(88) = -2.462$, $p < .05$, $r = .25$; Movies: $t(92) = -1.804$, $p > .05$, $r = .18$; Sales: $t(92) = -1.807$, $p > .05$, $r = .19$; Nazi: $t(90) = -5.15$, $p < .0001$, $r = .48$.

[8] $F(1, 368) = 31.127$, $p < .001$, $r = .28$.

Importantly, we found no main effect for gender in any of the four pairs of cases used in Experiment 1 and no interaction effects between gender and any of the experimental conditions.[9] Again, then, we have another failure to replicate Starmans and Friedman's (2009) initial results concerning gender differences in epistemic intuitions about Gettier cases.

The vignette pairs used by Beebe and Jensen (2012) differ from MOVIES, SALES and NAZI only by having elements of epistemic luck added to them. The vignette pair used by Beebe and Buckwalter (2010) differs in a few additional respects from WATER. Despite these differences and the fact that Beebe and Buckwalter (2010) and Beebe and Jensen (2012) used Likert scales that ranged from -3 to 3 instead of from 1 to 7, it can nevertheless be instructive to compare the mean responses obtained in these previous studies with those obtained in Experiment 1. The first row of Table 3 displays the mean responses from the experiments of Beebe and Buckwalter (2010) and Beebe and Jensen (2012) transformed to a 1 to 7 scale. The second row shows the mean responses from Experiment 1. Although it must be kept in mind that this kind of comparison has important limitations—because logically equivalent scales are often not pragmatically equivalent—what we appear to find is that Gettierizing factors drive down knowledge attributions but not as much as traditional philosophical wisdom would enjoin.[10]

	ENVIRONMENT/ WATER		**MOVIES**		**SALES**		**NAZI**	
	Help/Rise	**Harm/Fall**	**Improve**	**Drop**	**Increase**	**Decrease**	**Fulfill**	**Violate**
unGettiered	4.91	6.25	4.92	5.5	4.8	5.5	4.96	5.81
Gettiered	3.43	4.35	3.72	4.33	3.37	4.04	3.34	5.15

Table 3. Mean participant responses in the unGettiered cases used by Beebe and Jensen (2012) and the Gettiered cases of Experiment 1.

[9] Water (gender): $F(1, 89) = .976$, $p > .05$; Water (gender * condition): $F(1, 89) = 2.003$, $p > .05$; Movies (gender): $F(1, 86) = .432$, $p > .05$; Movies (gender * condition): $F(1, 86) = .728$, $p > .05$; Sales (gender): $F(1, 85) = .528$, $p > .05$; Sales (gender * condition) $F(1, 85) = .602$, $p > .05$; Nazi (gender): , $F(1, 88) = 1.398$, $p > .05$; Nazi (gender * condition): $F(1, 88) = .003$, $p > .05$.

[10] Cf. Cullen (2010) and Beebe and Jensen (forthcoming, sec. 2) for discussion of how using different scales can changes the pragmatics of an experimental situation.

Many philosophers (both mainstream and experimental) tend to make overly simplistic inferences from what a philosophical theory says about *X* to what that theory would predict regarding folk responses to *X*. However, as Keith DeRose (2011) and others have emphasized, a well developed philosophical theory about the ordinary concept of knowledge may nonetheless fail to be sufficiently well developed to make any prediction about how untrained participants will handle the concept of knowledge in experimental settings. Thus, while we do not want to oversimplify the relationship between the overwhelming consensus within professional philosophy over the last fifty years regarding Gettier cases, on the one hand, and psychological predictions about folk epistemic assessments, on the other, we nevertheless believe that many philosophers would find the results of Experiment 1 to be remarkable.

3. Experiment 2

In a further investigation of participants' intuitions about cases that combine elements from both the Gettier case and the Knobe effect literatures, we took several well-known Gettier cases and 'Knobified' them. In other words, we took protagonists from familiar Gettier cases and in some conditions had those protagonists engage in acts of moral turpitude. For each Gettier case and Knobified Gettier case pair, we also constructed an unGettiered counterpart—i.e., a case that lacks the kind of epistemic luck characteristic of Gettier cases—in order to have a control to which to compare the other two cases. Our prediction was that participants would be more likely to attribute knowledge in each of the Knobified Gettier cases than in either the original Gettier cases or the unGettiered controls. As we will see below, our predictions were half right and half wrong.

The first case we chose was the following one from Bryan Skyrms (1967), which in contrast to each of Gettier's original cases does not have the protagonist drawing any sort of inference to the Gettierized belief in question:

> MATCH1: A pyromaniac has just purchased a box of Sure-Fire Matches. He has done so many times before and has noted that they have always lit when struck unless they were wet. Furthermore, he knows that oxygen must be present for things to burn and that the observed regularity between matches' being struck and their lighting is not a mere coincidence. After perceiving that the matches are dry and that there is plenty of oxygen present, he proceeds to strike one of the matches, confident that it will light. It does. Unbeknownst to the pyromaniac, however, the match happens to contain impurities that prevent it from lighting simply by being struck. What ignited the match was an extremely rare burst of cosmic radiation that happened to arrive at just the right place at the very moment the match was being struck.

Participants were then asked the extent to which they agreed or disagreed with the following claim: "The pyromaniac knew that the match would light." Participants were then asked to select 'Strongly Disagree,' 'Disagree,' 'Neutral,' 'Agree,' or 'Strongly Agree' as their answer. A different answer format was used in Experiment 2 than in Experiment 1 because participants in Experiment 2 completed an online questionnaire rather than a pencil and paper one and having five rather than seven answer choices seemed better suited to the online platform that was used.

The following case was constructed in which many of the essential features of MATCH1 were preserved but where significant harm resulted from the match being lit:

> MATCH2: A deeply disturbed criminal has just purchased a box of Sure-Fire Matches. He wants to burn down a local orphanage in the middle of the night, killing all of the

children inside. He has purchased Sure-Fire Matches many times before and has noted that they have always lit when struck unless they were wet. Furthermore, he knows that oxygen must be present for things to burn and that the observed regularity between matches' being struck and their lighting is not a mere coincidence. During the middle of the night the criminal sneaks over to the orphanage and pours a large amount of highly flammable liquid all around the outside of the building. After perceiving that the matches are dry and that there is plenty of oxygen present, he proceeds to strike one of the matches, confident that it will light. It does. Unbeknownst to the criminal, however, the match happens to contain impurities that prevent it from lighting simply by being struck. What ignited the match was an extremely rare burst of cosmic radiation that happened to arrive at just the right place at the very moment the match was being struck.

Participants were then asked to indicate the extent to which they agreed or disagreed that the criminal knew that the match would light. Answer choices were the same as in MATCH1.

The following unGettiered case was also constructed, in which the elements of epistemic luck found in MATCH1 were omitted:

MATCH3: A pyromaniac has just purchased a box of Trusty Matches. He has done so many times before and has noted that they have always lit when struck unless they were wet. Furthermore, he knows that oxygen must be present for things to burn and that the observed regularity between matches' being struck and their lighting is not a mere coincidence. After perceiving that the matches are dry and that there is plenty of oxygen present, he proceeds to strike one of the matches, confident that it will light. It does.

Participants were asked to judge the extent to which they agreed or disagreed with the following claim: “The pyromaniac knew that the match would light.”

In a between subjects design 192 undergraduate college students (mean age = 27, 53% female, 73% Anglo-American) from a large, public university in the northeastern United States received one of the three match vignettes described above. The mean response for each vignette is represented in Figure 2. In each of the figures in this section, 'Strongly Disagree' is represented on the *y* axis as '1,' 'Disagree' as '2,' 'Neutral' as '3,' 'Agree' as '4,' and 'Strongly Agree' as '5.' As predicted, attributions of knowledge were highest in the significant harm version of MATCH. However, while the difference between the mean responses in MATCH2 and MATCH1 was statistically significant, the difference between the means in MATCH2 and MATCH3 was not.[11] Unsurprisingly, the difference between the average response in the Gettiered and unGettiered conditions (i.e., MATCH1 and MATCH3) was statistically significant as well. Strikingly, however, the most common response in each condition of the three conditions was 'Agree.' Moreover, 53% of participants in MATCH1 chose either 'Agree' or 'Strongly Agree,' while 77% in MATCH2 and 78% in MATCH3 did so as well.

[11] A one-way ANOVA reveals a significant effect for vignette type ($F(2, 171) = 12.470$, $p < .001$, $r = .3$). A post-hoc Tukey's HSD test revealed a significant difference (at the 0.01 level) between the mean response in MATCH1 and the mean responses in the other two conditions but no significant difference ($p > .05$) between the MATCH2 and MATCH3 conditions. The test also revealed a significant difference (at the 0.01 level) between the means in MATCH1 and MATCH3.

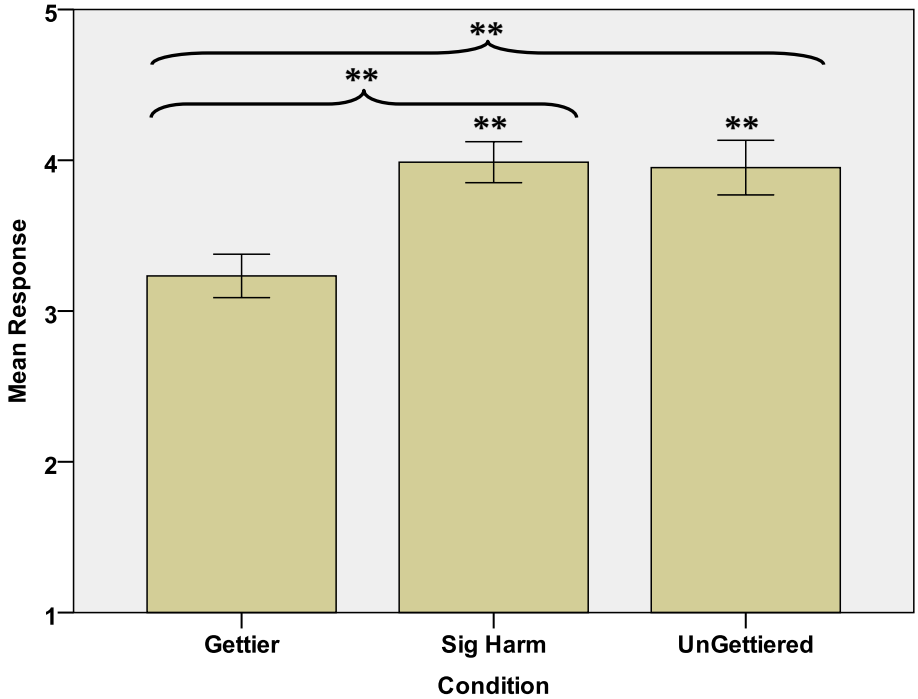

Figure 2. Mean participant responses in the Gettier (*M* = 3.23, *Md* = 4, *SE* = .14), significant harm (*M* = 3.99, *Md* = 4, *SE* = .14) and unGettiered (*M* = 3.95, *Md* = 4, *SE* = .18) conditions of the MATCH vignette set.

Importantly, the mean response in MATCH2—a Gettier case—was significantly above the midpoint with a large effect size.[12] The mean response in MATCH1 did not differ significantly from the midpoint, but this is itself a noteworthy result, since MATCH1 is Skyrms' original Gettier case.[13] From the perspective of contemporary epistemology, the 78% who ascribed knowledge to the pyromaniac in MATCH3 seem to be on track, but the knowledge ascribers in the other two conditions seem to have gone off the rails. Similar results were found with three other sets of cases.

A second set of Knobified Gettier cases was constructed around the following widely discussed case due to Gilbert Harman (1973, p. 143):

[12] $t(77) = 7.262$, $p < .001$, $r = .64$. The mean in MATCH3 was also significantly above the midpoint: $t(40) = 5.247$, $p < .001$, $r = .64$.

[13] $t(72) = 1.617$, $p > .05$, $r = .19$.

MAIL1: Gilbert's friend Donald tells Gilbert that he's going to Italy for the summer. In June Gilbert takes Donald to the airport and see him off. In July Donald decides to send Gilbert several letters informing him that he has gone to San Francisco. This is not true. Donald is simply trying to fool Gilbert. Donald sends the letters to another friend in San Francisco who is instructed to send them to Gilbert one at a time, as if they were sent from Donald, complete with a San Francisco postmark. Gilbert has not read any of these letters because he has been out of town. When he returns home, Gilbert's mail has piled up. Standing before a pile of unopened mail that includes two of Donald's phony letters, Gilbert still believes that Donald is in Italy. He is right. Donald is in Italy.

Philosophers who draw a distinction between Gettier cases and barn-façade cases would likely count MAIL1—and, as we will see below, POLITICIAN1—as instances of the latter rather than the former. However, since MAIL1 involves the kind of double epistemic luck that for present purposes we are taking to characterize Gettier cases, we will treat it as one. Participants were asked the extent to which they agreed or disagreed that Gilbert knows that Donald is in Italy and were again given the answer choices of 'Strongly Disagree, 'Disagree,' 'Neutral,' 'Agree,' and 'Strongly Agree.'

In Experiment 1, building upon the work of Beebe and Buckwalter (2010) and Beebe and Jensen (2012), we found that wrongdoers were more likely to be viewed as knowers than their good or neutral counterparts. Although some experimental philosophers (e.g., Schaffer and Knobe 2012) have suggested that simply having wrongdoing in a vignette might make participants more likely to attribute knowledge to characters in the story who were simply innocent or neutral bystanders, we hypothesized that it would not have this kind of effect. In order to investigate the issue, we transformed Donald—the seemingly innocent trickster of

Harman's original case—into something more sinister in the following significant harm case, and then we asked participants whether Gilbert—the innocent bystander—had knowledge.[14]

> MAIL2: Donald is hiding in the bushes behind a house in a quiet neighborhood, waiting for 8 year old Amelia to walk home from school. As she rounds the corner and comes into sight, Donald grabs her, clamps his hand down tight over her mouth, and stuffs her into his nearby, waiting van. Donald then flees the country on a plane bound for Italy, with his hostage in tow. Donald tells his friend Gilbert that he is going to Italy for a summer vacation. In July, in order to throw law enforcement officials off his track, Donald decides to send Gilbert several letters informing him that he has gone to San Francisco. This is not true. Donald continues to stay in Italy. Donald sends the letters to another friend in San Francisco who is instructed to send them to Gilbert one at a time, as if they were sent from Donald, complete with a San Francisco postmark. Gilbert has not read any of these letters because he has been out of town. When he returns home, his mail has piled up. Standing before a pile of unopened mail that includes two of Donald's phony letters, Gilbert still believes that Donald is in Italy. He is right. Donald is in Italy.

Participants were asked whether they agreed that Gilbert knows that Donald is in Italy. Finally, the following unGettiered version of the original vignette was constructed for the sake of comparison:

> MAIL3: Gilbert's friend Donald tells Gilbert that he's going to Italy for the summer. In June Gilbert takes Donald to the airport and sees him off. During July Donald sends Gilbert several letters that are postmarked from Italy. Gilbert believes that Donald is in Italy. He is right. Donald is in Italy.

Participants were asked the same question as in the other two cases.

[14] Thanks to DS for helping us clarify our explanation of this case and its experimental upshot.

In a between subjects design 221 undergraduate college students (mean age = 30, 53% female, 83% Anglo-American) from a large, public university in the northeastern United States were given one of the MAIL vignettes above. Their mean responses are represented in Figure 3. Participants were not more likely to ascribe knowledge in the significant harm version of MAIL (i.e., MAIL2), with almost half of them (42%) choosing 'Neutral.' Nevertheless, this is a significant result, since MAIL2 is a Gettier case. Participants were much more likely to ascribe knowledge in the unGettiered MAIL3 than in the other conditions, but a striking 53% of participants in the basic Gettier case (i.e., MAIL1) either agreed or strongly agreed that Gilbert knows Donald is in Italy, whereas only 30% disagreed or strongly disagreed.[15] Although the mean participant response in MAIL2 was not significantly above the neutral midpoint, the mean in MAIL1 was.[16] Thus, we again see that folk attributions of knowledge in Gettier cases are higher than many philosophers might have expected.

[15] A one-way ANOVA reveals a significant effect for vignette type ($F(2, 218) = 29.462$, $p < .001$, $r = .52$). A post-hoc Tukey's HSD test revealed a significant difference (at the 0.01 level) between the mean response in the MAIL3 condition and the mean responses in the other two conditions but no significant difference ($p > .05$) between the MAIL1 and MAIL2 conditions. The effect size for condition was large.

[16] MAIL1: $t(100) = 2.348$, $p < .05$, $r = .23$. MAIL2: $t(59) = 1.026$, $p > .05$, $r = .13$. MAIL3: $t(59) = 19.55$, $p < .001$, $r = .93$.

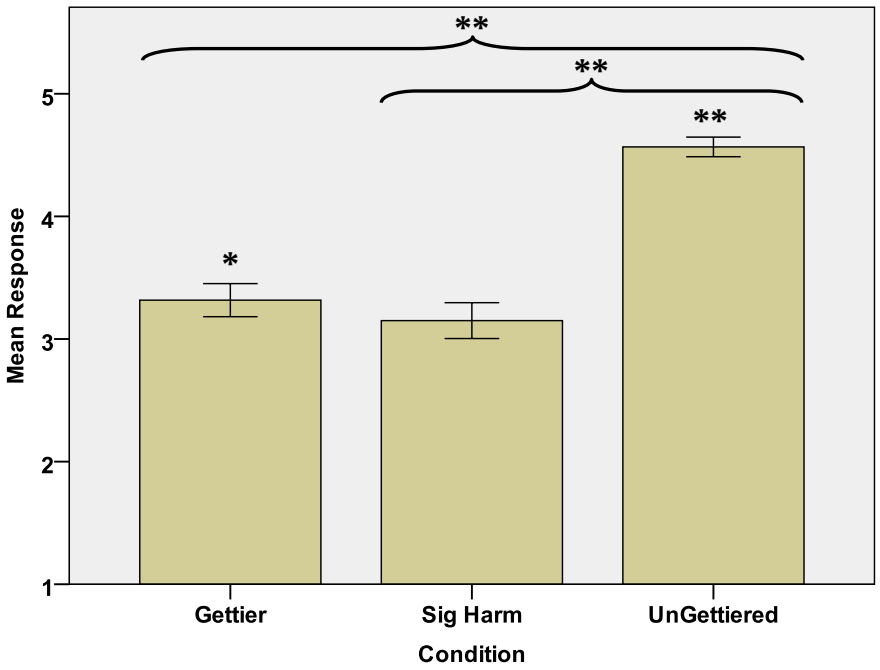

Figure 3. Mean participant responses in the Gettier (*M* = 3.32, *Md* = 4, *Mo* = 4, *SE* = .14), significant harm (*M* = 3.15, *Md* = 3, *Mo* = 3, *SE* = .15) and unGettiered (*M* = 4.57, *Md* = 5, *Mo* = 5, *SE* = .08) conditions of the MAIL vignette set.

A third set of Knobified Gettier cases also drew inspiration from the influential work of Harman (1973, pp. 143-144), the first member of which was the following:

POLITICIAN1: A political leader is assassinated. His associates, fearing a coup, decide to pretend that the bullet hit someone else. On nationwide television they announce that an assassination attempt has failed to kill the leader but has killed a secret service agent by mistake. However, before the announcement is made, an enterprising reporter on the scene faxes the real story to her news agency so that the story can be included in the day's final edition of the paper. Jill buys a copy of that paper and reads the story of the assassination that was dictated by the reporter who witnessed the event. Unlike most everyone else, Jill has not heard about the false television report.

Participants were asked "Please indicate the extent to which you agree or disagree with the following claim: 'Jill knows that the political leader has been assassinated.'" Answer choices again ranged from 'Strongly Disagree' to 'Strongly Agree.' Although elements of wrongdoing were already present in POLITICIAN1—an assassination, a coup, a cover up—the putative knower was not herself a wrongdoer. She was merely a neutral observer. In the following variation on Harman's original case, the putative knower becomes the central doer of wrong:

> POLITICIAN2: Ivan plans to assassinate a recently elected politician. Because the politician's bodyguards and associates fear for his life, they decide to pretend that the politician is away from the capital city until they can formulate a better plan for his protection. On nationwide television they announce that the politician is taking a two-week journey to another country. However, before the announcement is made, a reporter closely following the politician faxes a story containing correct information about the politician's whereabouts to her news agency, and the story is included in the day's final edition of the paper. Ivan buys a copy of that paper and reads the correct information about the politician's location. Unlike most everyone else, Ivan has not heard about the false television report. Ivan then travels to the location specified in the newspaper and assassinates the politician and his bodyguards.

Participants were asked the extent to which they agreed or disagreed that Ivan knew where the politician could be found.[17] The following, unGettiered version of the original was also constructed for the sake of comparison:

[17] KM has raised the following concern about the wording of POLITICIAN2. The vignette does not make it as clear as it should that in thinking about whether Ivan knew where the politician could be found participants were supposed to be thinking of whether Ivan had knowledge after reading the newspaper but before seeing and assassinating the politician. While we agree that the vignette does not foreclose this possibility as much as it could have, we believe the most charitable way for participants to interpret the vignette is in accord with its intended meaning.

POLITICIAN3: A political leader is assassinated. A reporter on the scene sends news of the assassination to her news agency so that the story can be included in the day's final edition of the paper. Jill buys a copy of that paper and reads the story of the assassination that was dictated by the reporter who witnessed the event.

Participants were asked they agreed that Jill knows that the political leader has been assassinated.

In a between subjects design 189 undergraduate college students (mean age = 31, 53% female, 79% Anglo-American) from a large, public university in the northeastern United States were given one of the POLITICIAN cases above. Their mean responses are represented in Figure 4. As expected, participants given the significant harm version (i.e., POLITICIAN2) were more likely to attribute knowledge to the central protagonist than in Harman's original POLITICIAN case. These results provide further confirmation for the hypothesis that participants are not equally likely to attribute knowledge to observers of wrongdoing as they are to attribute knowledge to doers of wrong. Participants were not, however, more likely to attribute knowledge in POLITICIAN2 than in POLITICIAN3, despite the fact that 83% of participants in the former condition agreed or strongly agreed that Ivan knows that the political leader has been assassinated, while only 65% in the latter gave a similar verdict about Jill.[18] Importantly, the mean responses in all three conditions of the POLITICIAN vignette set were significantly above the neutral midpoint, with medium to extremely large effect sizes.[19]

[18] A one-way ANOVA reveals a significant effect for vignette type ($F(2, 186) = 10.529$, $p < .01$, $r = .35$). A post-hoc Tukey's HSD test revealed a significant difference (at the 0.01 level) between the mean response in the POLITICIAN1 condition and the mean responses in the other two conditions but no significant difference ($p > .05$) between the other two conditions. The effect size for condition was medium.

[19] POLITICIAN1: $t(68) = 3.869$, $p < .001$, $r = .42$. POLITICIAN2: $t(59) = 8.603$, $p < .001$, $r = .75$. POLITICIAN3: $t(59) = 12.157$, $p < .001$, $r = .85$.

Figure 4. Mean participant responses in the Gettier (*M* = 3.55, *Md* = 4, *Mo* = 4, *SE* = .14), significant harm (*M* = 4.18, *Md* = 4, *Mo* = 5, *SE* = .14) and unGettiered (*M* = 4.35, *Md* = 5, *Mo* = 5, *SE* = .11) conditions of the POLITICIAN vignette set.

The final set of Knobified Gettier cases began with the following vignette from Simon Cullen (2010), which was inspired by Bertrand Russell's (1912) famous stopped clock case:

> CLOCK1: Mary works as a clerk in an office. She is clear-headed and has excellent eye-sight. Mary knows that she set the clock above her desk accurately and that it has been completely reliable for many years. At 3:00pm Mary looks up at the clock and sees that it reads "3:00pm," and indeed, it is 3:00pm. However, unknown to Mary, the clock stopped working exactly 24 hours ago.

Participants were asked whether Mary knows that the time is 3:00pm. The following significant harm version was constructed by modifying CLOCK1:

> CLOCK2: John is a terrorist. He plans to detonate a bomb in a crowded city shopping mall at a time when a local politician will be arriving to give a press conference there. John is clear-headed and has excellent eye-sight. He knows that he set the time of his

wristwatch accurately and that it has been completely reliable for many years. At 3:00pm John looks at his watch and sees that it reads "3:00pm," and indeed, it is 3:00pm. However, unbeknownst to John, his watch stopped working exactly 24 hours ago. John detonates the bomb, killing the politician and dozens of bystanders.

Finally, an unGettiered version of CLOCK1 was constructed for the sake of comparison:

CLOCK3: Wendy works as a clerk in an office. She is clear-headed and has excellent eye-sight. Wendy knows that she set the clock above her desk accurately and that it has been completely reliable for many years. At 3:00pm Wendy looks up at the clock and sees that it reads "3:00pm," and indeed, it is 3:00pm.

In a between subjects design 175 undergraduate college students (mean age = 26, 53% female, 85% Anglo-American) from a large, public university in the northeastern United States were given one of the CLOCK cases above. Mean responses are represented in Figure 5. The mean response in CLOCK2 fell in between the means for CLOCK1 and CLOCK3, and the difference between the CLOCK2 mean and the means in the other two conditions failed to be statistically significant.[20] However, the difference between the CLOCK1 and CLOCK3 means was statistically significant. Only the mean for CLOCK3 was significantly above the midpoint.[21]

[20] A one-way ANOVA reveals a significant effect for vignette type ($F(2, 172) = 5.222$, $p < .01$, $r = .25$). A post-hoc Tukey's HSD test revealed a significant difference (at the 0.01 level) between the mean response in the CLOCK1 and CLOCK3 conditions but no significant difference ($p > .05$) between the CLOCK2 and the other two conditions. The effect size for condition was small.

[21] CLOCK1: $t(74) = .65$, $p > .05$, $r = .08$. CLOCK2: $t(59) = 1.918$, $p > .05$, $r = .24$. CLOCK3: $t(39) = 4.892$, $p < .001$, $r = .62$.

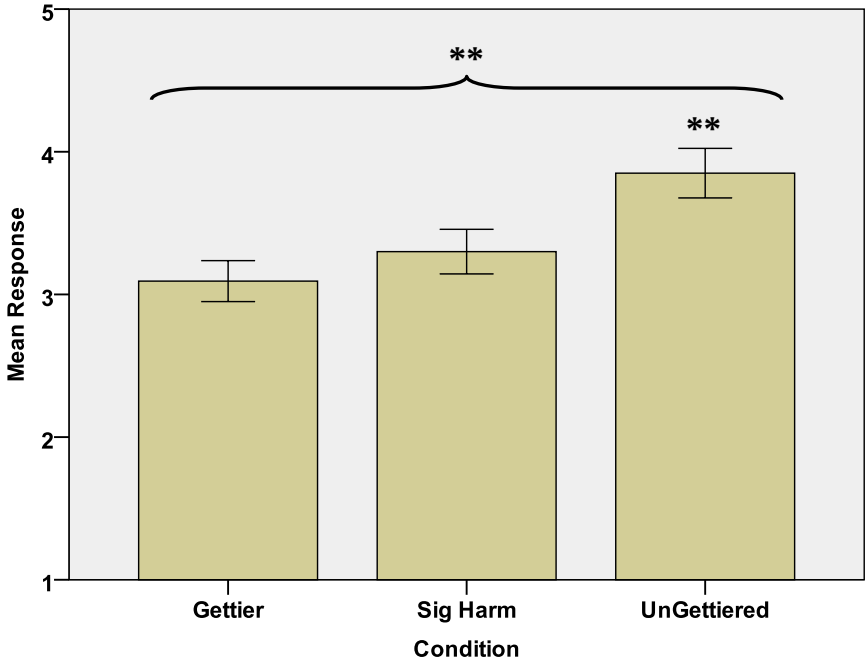

Figure 5. Mean participant responses in the Gettier (*M* = 3.09, *Md* = 3, *Mo* = 4, *SE* = .14), significant harm (*M* = 3.30, *Md* = 4, *Mo* = 4, *SE* = .16) and unGettiered (*M* = 3.85, *Md* = 4, *Mo* = 4, *SE* = .17) conditions of the CLOCK vignette set.[22]

Collapsing participant responses across all four sets of cases used in Experiment 2 yields the comparison represented in Figure 6. Participants were on the whole more likely to attribute knowledge in every case than they were to deny knowledge to the central protagonists. All of the collapsed means fell significantly above the midpoint and differed significantly from one another at the .01 level.[23] The 'Knobifying' effect was also significant, although not quite as large as we had originally predicted. In none of the four sets of cases used in Experiment 2 was there a main effect for gender or an interaction effect between gender and any of the separate conditions.[24]

[22] Using a forced-choice format, Nagel (forthcoming b) reports that 44% of participants she presented with a stopped clock Gettier case ascribed knowledge to the protagonist of the story.

[23] Gettier: $t(317) = 4.155$, $p < .001$, $r = .23$. Significant harm: $t(239) = 9.423$, $p < .001$, $r = .52$. UnGettiered: $t(200) = 18.237$, $p < .001$, $r = .79$. A one-way ANOVA reveals a significant effect for vignette type ($F(2, 756) = 40.778$, $p < .001$, $r = .33$). A post-hoc Tukey's HSD test revealed a significant difference (at the 0.001 level) between each condition and the other two conditions. The effect size for condition was medium.

[24] MATCH (gender): $F(1, 183) = .849$, $p > .05$; MATCH (gender * condition): $F(2, 183) = 1.281$, $p > .05$; MAIL (gender): $F(1, 221) = .965$; MAIL (gender * condition): $F(2, 221) = .201$, $p > .05$; POLITICIAN (gender): $F(1, 189) = .108$, $p > .05$; POLITICIAN (gender * condition): $F(2, 189) = 1.432$, $p > .05$; CLOCK (gender): $F(1,$

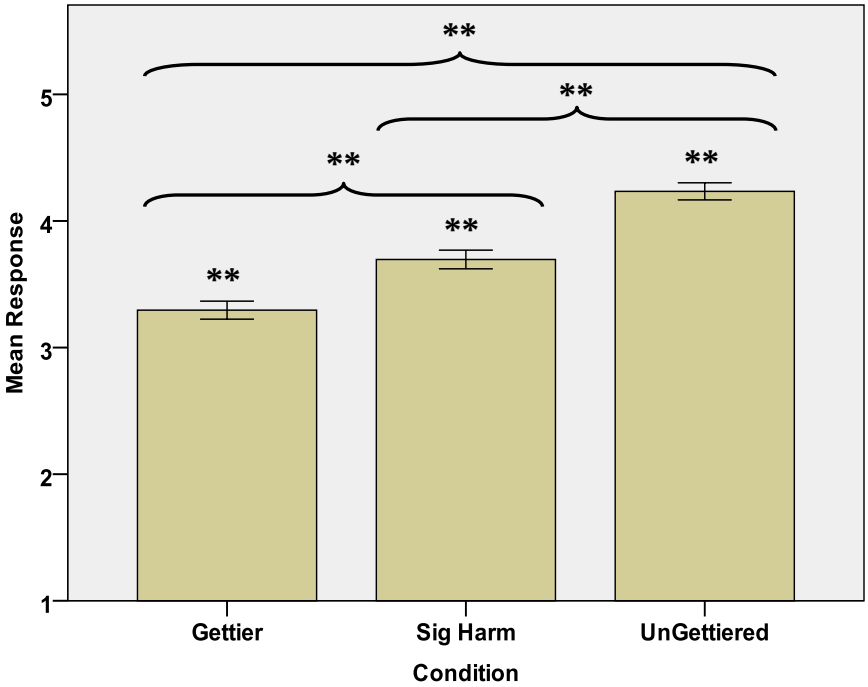

Figure 6. Mean participant responses in the collapsed Gettier (*M* = 3.30, *Md* = 4, *Mo* = 4, *SE* = .07), significant harm (*M* = 3.70, *Md* = 4, *Mo* = 4, *SE* = .07) and unGettiered (*M* = 4.23, *Md* = 4, *Mo* = 5, *SE* = .07) conditions.

4. Conclusion

Too many experimental philosophers make the mistake of assuming that psychological theses about how ordinary participants will respond to experimental materials follow quite directly and unproblematically from philosophical theses about the nature of knowledge or about the correct way to view those materials. Philosophers who believe that the correct judgment to make about a particular thought experiment is that the protagonist fails to have knowledge might think that most ordinary people will agree with this verdict. But such a belief is not required by the mere fact that these philosophers believe this is the correct way to view the thought experiment. They

175) = .030, $p > .05$; CLOCK (gender * condition): $F(2, 175) = .947$, $p > .05$.

might simply think that most other philosophers would agree with them or that anyone who correctly apprehends certain independent theoretical considerations would agree. Such a view leaves open the possibility that those who lack formal philosophical training or fail to grasp the relevant theoretical considerations might very well not make the same epistemic assessment.

Nonetheless, while we do not want to oversimplify the relationship between philosophical theses and psychological predictions, we do believe that running throughout the Gettier literature has been a fairly strong (if unspoken) consensus about how ordinary subjects would in fact respond to those cases. According to this consensus, in the absence of formal training in philosophy, the apprehension of abstract principles about the nature of knowledge, or even much intellectual acumen, most of the folk should judge that protagonists in Gettier cases fail to have knowledge. We take this psychological thesis to be challenged by our results.

The two most pressing explanatory questions raised by our results concern the reasons why participants attributed knowledge as often or as strongly as they did in Gettier cases and why they were more inclined to attribute knowledge in 'Knobified' cases than in their 'unKnobified' counterparts. Starmans and Friedman (2012) recently found that participants were more willing to say that protagonists in several Gettier cases 'really know' certain facts than they were to say the protagonists 'only believe' or 'only think' the facts are true. According to Starmans and Friedman (2012, p. 279), their findings suggest that "people's folk conception of knowledge fits, at least roughly, with the classical philosophical view of knowledge as justified true belief, and conflicts with the current philosophical conception in which Gettiered individuals are viewed as not possessing knowledge." While we find this suggestion to represent an intriguing possibility, we think it is a bit premature, since participants in other studies have shown a reluctance to attribute knowledge in Gettier cases.

For example, in a study of eight different Gettier cases, Nagel et al. (forthcoming) report that an average of 33% of participants ascribed knowledge to the central protagonists. This figure appears to fall significantly below the midpoint of 50%. By contrast, in the sixteen Gettier cases reported above (eight each in Experiments 1 and 2), mean participant responses fell significantly below the midpoint three times, significantly above the midpoint five times, and did not differ significantly from the midpoint eight times. One reason for the apparent discrepancy between the findings of Nagel et al. and the findings that we report may be that Nagel et al. used a forced-choice format that asked participants to choose between two non-exhaustive options—either '[the protagonist] knows that p' or '[the protagonist] thinks that p, but doesn't know that p'—whereas we employed Likert scales that gave participants a range of responses. Participants' responses have been shown to differ in surprising ways between forced-choice and Likert scale formats and between exhaustive and non-exhaustive forced choice formats, even when identical research materials are used.

Nagel (forthcoming a) suggests that low levels of participant motivation might be an important factor in keeping some participants from denying knowledge in Gettier cases:

> Participants who are not interested in a particular story may be more inclined to respond to it randomly. Philosophers and others may have the same basic intuitive capacity to register the presence or absence of knowledge, but philosophers may be more motivated to read epistemic vignettes with an eye to exercising this capacity.

Responding randomly to research materials pushes the distribution of answers toward the midpoint and away from either extreme. The traditionally correct response to Gettier cases, of course, lies at one extreme. One important factor that gives philosophers an increased level of motivation when reading such stories is their appreciation of the broader dialectical context in

which the thought experiment is offered as an intuitive test of an epistemological theory.

Relatedly, throughout the data obtained by experimental epistemologists participants appears to display a general bias in favor of agreement, as opposed to disagreement. Since most experimental materials ask participants whether protagonists know but rarely ask if protagonists fail to know, the data that has been obtained so far may not fully reflect participants' true epistemic intuitions, if those intuitions are filtered through a general tendency to be agreeable or accommodating.[25] For this reason, we have not placed as much emphasis on the fact that participants were found to attribute knowledge in Gettier cases as we have on the fact that participants were more inclined to attribute knowledge in Gettier cases when wrongdoing was involved.

Regarding the question of why participants in our studies were more inclined to attribute knowledge in 'Knobified' cases than in their 'unKnobified' counterparts, Alfano, Beebe, and Robinson (2012) hypothesize that having practical reasons for paying attention to certain possibilities (a) makes rational agents more likely to form beliefs (and more likely to form stronger beliefs) about those possibilities than about possibilities they have less practical reason to consider and (b) makes rational observers more likely to attribute beliefs (and greater degrees of belief) about those possibilities to agents. Although knowledge is not equivalent to firmly held belief, Alfano, Beebe, and Robinson hypothesize that an increased willingness to attribute belief or stronger degrees of belief often leads ordinary participants to be more likely to attribute knowledge as well. Describing the thought process of the protagonist in Knobe's original environment case, Alfano, Beebe, and Robinson (2012, p. 269) write:

> The chairman in the HELP condition, for example, does not need to say to himself, "Wait! I need to stop and think carefully about whether helping the environment is

[25] Thanks to KM for getting us to see the relevance of this issue.

> something that I should be doing." In the HARM condition, however, an inner monologue like this might well be appropriate. The same seems to hold for the CEO who is considering violating or fulfilling a racial identification law in Nazi Germany and indeed for any of the other protagonists in the Knobe effect literature.

Because of the potential costs involved, the chairman who harms the environment, the movie studio executive who decreases the quality of his movies, and the CEO who violates a Nazi law all need to consider with some degree of care whether they are embarking upon the right course of action, whereas their counterparts do not seem to have a similar need (or at least do not have as significant a need to do so). Alfano, Beebe, and Robinson suggest that the need to engage in careful reflection leads participants to think that protagonists in harmful or blameworthy conditions are more likely to have knowledge about the outcomes of their actions than protagonists in good or neutral conditions. They also suggest that the stronger tendency to attribute belief in harm cases than in help cases may lead participants to be more likely to attribute intentionality in the former kind of case.

The central virtue of Alfano, Beebe, and Robinson's account is that it provides an elegant explanation of the original Knobe effect, the epistemic side-effect effect, the Knobe effect for belief ascriptions (Beebe forthcoming), and the data reported above. However, the fact remains that something like the Knobe effect has been found across a dizzying variety of domains and that (despite the confident proclamations of some) the experimental philosophy community has no very clear idea about whether or how all of these cases might be explained by a single, overarching theory. For this reason, while we think the suggestion of Alfano, Beebe, and Robinson holds some promise, we do not think it would be wise at this stage of the empirical

investigation of these folk psychological notions to invest in it a markedly high degree of confidence.

Another issue that should be examined in further detail is whether more reflective participants or those who have been primed to engage in deeper levels of reflection will be as likely as our participants to attribute knowledge in Gettier and Knobe effect cases. Using Shane Frederick's (2005) cognitive reflection test—a measure of reflectiveness that is strongly correlated with general intelligence—N. Ángel Pinillos et al. (2011) found that greater reflectiveness was correlated with a decreased asymmetry in attributions of intentionality in the help and harm versions of Knobe's original ENVIRONMENT case. In other words, there was less of a Knobe effect among the more reflective. It would be interesting to see the extent to which greater reflectiveness is associated with decreased asymmetry in knowledge attributions across a variety of Knobe effect cases and what effect greater reflectiveness might have on participants' willingness to attribute knowledge in Gettier cases.

In conclusion, despite the fact that it seems too early to know how to explain the full range of Knobe effects, our results demonstrate some further respects in which folk epistemic assessments can be influenced by what appear to be epistemically irrelevant factors. If Gettierizing factors are capable of robbing a subject of knowledge, it seems they should do so regardless of whether that subject is performing positive or negative actions. The fact that folk epistemic assessments are affected by such factors seems to provide an additional reason why untutored epistemic intuitions might be a poor foundation upon which to base a philosophical account of the true nature of knowledge.

References

Alfano, Mark, Beebe, James R., and Robinson, Brian. 2012. 'The Centrality of Belief and Reflection in Knobe Effect Cases: A Unified Account of the Data.' *The Monist*, 95: 264-89.

Alicke, Mark D. 2008. 'Blaming Badly.' *Journal of Cognition and Culture*, 8 179-86.

Beebe, James R. forthcoming. 'A Knobe Effect for Belief Ascriptions.' *The Review of Philosophy and Psychology*.

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

Beebe, James R., and Jensen, Mark. 2012. 'Surprising Connections Between Knowledge and Intentional Action: The Robustness of the Epistemic Side-Effect Effect.; *Philosophical Psychology* 25: 689-715.

Buckwalter, Wesley. forthcoming. 'Gettier Made ESEE.'

Cullen, Simon. 2010. 'Survey-Driven Romanticism.' *Review of Philosophy and Psychology*, 1: 275-96.

DeRose, Keith. 2011. 'Contextualism, Contrastivism, and X-Phi Surveys.' *Philosophical Studies*, 156: 81-110.

Frederick, Shane. 2005. 'Cognitive Reflection and Decision Making.' *The Journal of Economic Perspectives*, 19: 25-42.

Gettier, Edmund L. 1963. 'Is Justified True Belief Knowledge?' *Analysis*, 23: 121-3.

Harman, Gilbert. 1973. *Thought*. Princeton: Princeton University Press.

Hetherington, Stephen Cade. 1999. 'Knowing Failably.' *Journal of Philosophy*, 96: 565-87.

Knobe, Joshua. 2003a. ‘Intentional Action and Side Effects in Ordinary Language.’ *Analysis*, 63: 190-3.

_____. 2003b. ‘Intentional Action in Folk Psychology: An Experimental Investigation.’ *Philosophical Psychology*, 16: 309-24.

_____. 2004. ‘Folk Psychology and Folk Morality: Response to Critics.’ *Journal of Theoretical and Philosophical Psychology*, 24: 270-9.

_____. 2007.

Knobe, Joshua, and Mendlow, Gabriel. 2004. ‘The Good, the Bad, and the Blameworthy: Understanding the Role of Evaluative Considerations in Folk Psychology.’ *Journal of Theoretical and Philosophical Psychology*, 24: 252-8.

Koriat, Asher. 2008. ‘Subjective Confidence in One’s Answers: The Consensuality Principle.’ *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 34: 945-59.

_____. 2011. ‘Subjective Confidence in Perceptual Judgments: A Test of the Self-Consistency Model.’ *Journal of Experimental Psychology: General*, 140: 117-39.

Lewis, David. 1983. ‘Introduction.’ *Philosophical Papers, Vol. 1*. Oxford: Oxford University Press, pp. xx-xx.

Nadelhoffer, Thomas. 2004. ‘On Praise, Side Effects, and Folk Ascriptions of Intentionality.’ *Journal of Theoretical and Philosophical Psychology*, 24: 196-213.

Nagel, Jennifer. forthcoming a. ‘Intuitions and Experiments: A Defense of the Case Method.’ *Philosophy and Phenomenological Research*.

_____. forthcoming b. ‘Mindreading in Gettier Cases and Skeptical Pressure Cases.’ In Jessica Brown & Mikkel Gerken (eds.), *Knowledge Ascriptions*. Oxford: Oxford University Press.

Nagel, Jennifer, San Juan, Valerie, and Mar, Raymond. forthcoming. 'Gettier Case Recognition.'

Pinillos, N. Ángel, Smith, Nick, Nair, G. Shyam, Marchetto, Peter, and Mun, Cecilea. 2011. 'Philosophy's New Challenge: Experiments and Intentional Action.' *Mind & Language*, 26: 115-39.

Russell, Bertrand. 1912. *The Problems of Philosophy*. Home University Library.

Schaffer, Jonathan, and Knobe, Joshua. 2012. 'Contrastive Knowledge Surveyed.' *Noûs* 46: 675-708.

Skyrms, Bryan. 1967. 'The Explication of 'X Knows that P.'' *Journal of Philosophy*, 64: 373-89.

Starmans, Christina, and Friedman, Ori. 2009. 'Is Knowledge Subjective? A Sex Difference in Adults' Epistemic Intuitions.' Poster presented at the Biennial Meeting of the Cognitive Development Society, San Antonio, TX.

Starmans, Christina, and Friedman, Ori. 2012. 'The Folk Conception of Knowledge.' *Cognition* 124: 272-83.

Turri, John. 2011. 'Manifest Failure: The Gettier Problem Solved.' *Philosophers' Imprint*, 11: 1-11.

_____. 2012. 'Is Knowledge Justified True Belief?' *Synthese* 184: 247-59.

Weinberg, Jonathan M., Nichols, Shaun, and Stich, Stephen. 2001. 'Normativity and Epistemic Intuitions.' *Philosophical Topics* 29: 429-60.