

# A New Paradigm for Epistemology From Reliabilism to Abilism\*

John Turri

Philosophy Department and Cognitive Science Program  
University of Waterloo  
200 University Avenue West  
Waterloo, Ontario N2L3G1  
Canada  
[john.turri@gmail.com](mailto:john.turri@gmail.com)

*Abstract:* Contemporary philosophers nearly unanimously endorse *knowledge reliabilism*, the view that knowledge must be reliably produced. Leading reliabilists have suggested that reliabilism draws support from patterns in ordinary judgments and intuitions about knowledge, luck, reliability, and counterfactuals. That is, they have suggested a *proto-reliabilist hypothesis* about “commonsense” or “folk” epistemology. This paper reports nine experimental studies (N = 1262) that test the proto-reliabilist hypothesis by testing four of its principal implications. The main findings are that (a) commonsense fully embraces the possibility of unreliable knowledge, (b) knowledge judgments are surprisingly insensitive to information about reliability, (c) “anti-luck” intuitions about knowledge have nothing to do with reliability specifically, and (d) reliabilists have mischaracterized the intuitive counterfactual properties of knowledge and their relation to reliability. When combined with the weakness of existing arguments for reliabilism and the recent emergence of well supported alternative views that predict the widespread existence of unreliable knowledge, the present findings are the final exhibit in a conclusive case for abandoning reliabilism in epistemology. I introduce an alternative theory of knowledge, *abilism*, which outperforms reliabilism and well explains all the available evidence.

*Keywords:* knowledge; knowledge attributions; luck; counterfactuals; ability; abilism; reliabilism

---

\* This is the penultimate version of a paper to appear in *Ergo*. Please cite the final, published version if possible.

## Introduction

Contemporary philosophers have hotly debated the relationship between reliability and epistemic justification. Some argue that justification requires reliability (Goldman 1979, Swain 1981, Alston 1988, Greco 2013), while others roundly reject this (e.g. Lehrer & Cohen 1983, Cohen 1984, Foley 1985, Feldman 2013). By contrast, virtually every philosopher to address the matter accepts that knowledge requires reliability (e.g. Russell 1948, Sellars 1963, Dretske 1971, Armstrong 1973, Nozick 1981, Goldman 1986, Sosa 1991, Alston 1993, Plantinga 1993, Heller 1995, Zagzebski 1996, Lehrer 1997, Williamson 2000, Greco 2000, Bonjour 2002, Kornblith 2002, Riggs 2002, Cohen 2002, Feldman 2003, Roush 2005, Fumerton 2006, Pritchard 2009; Foley 1985 and Sartwell 1992 dissent). On this popular view, knowledge must be produced by abilities (processes, faculties, powers, etc.) that “will or would yield mostly true beliefs,” where an ability yields mostly true beliefs only if “considerably more than half” of the beliefs it produces are true (Alston 1995: 8-9; see also Vogel 2000; Comesaña 2002: 249; Goldman 2008; for a more extensive list of citations and quotations, see Turri 2015a). Call this consensus view *knowledge reliabilism*. Adapting a Quinean coinage, it’s not unfair to label knowledge reliabilism the central dogma of contemporary epistemology.

The literature contains surprisingly little explicit argumentation for knowledge reliabilism (hereafter just ‘reliabilism’). It is often just asserted, without elaboration, that knowledge requires reliability. The main passage cited in favor of reliabilism is a brief explanatory argument from an early paper by Alvin Goldman (1979), later dubbed the “master argument” lending relia-

bilism “powerful *prima facie* support” (Goldman 2012: 4). Goldman first notes that knowledge must be appropriately caused and then asks what kind of causes yield knowledge.<sup>1</sup> Goldman produces two lists. One list features processes that intuitively do produce knowledge: perception, introspection, memory, and “good reasoning.” The other list features processes that intuitively do not: “wishful thinking,” “mere hunch,” “guesswork,” and “confused reasoning.” Goldman notes that members of the first list all seem to be reliable, whereas members of the latter all seem to be unreliable. He then proposes an explanation for this: knowledge requires reliability.

But Goldman doesn’t consider alternative explanations. For example, the hypothesis that knowledge requires *ability*, reliable or not, can explain the membership of these lists. The list of processes that produce knowledge includes cognitive abilities to detect, discover, and retain truths; the list of processes that don’t produce knowledge doesn’t include abilities of detection, discovery, and retention. The weaker *abilist* hypothesis renders the *reliabilist* hypothesis otiose.

Not only is the best *prima facie* evidence for reliabilism equivocal, there are also positive arguments for the possibility of unreliable knowledge (Turri 2015a). One such argument begins with the observation that achievements can be unreliably produced. For example, human toddlers around twelve months old are highly unreliable walkers, but many of their early steps are genuine achievements that manifest their blossoming bipedalism. To take a less cuddly example,

---

1 Goldman’s discussion pertains to justification explicitly, but he also says that justification is necessary for knowledge (Goldman 1979: 340). Other writers reject the connection between justification and reliability but accept Goldman’s view about the relationship between knowledge and reliability (e.g. Feldman 2003: 90ff; Fumerton 2006: 86ff).

hockey players fail to score a goal on approximately ninety percent of their attempts, but many of their goals are genuine achievements that manifest impressive athletic ability. More generally, achievements populate the road to proficiency in many spheres, despite our unreliability. This is true in art, athletics, politics, oratory, music, science, and elsewhere. If achievement in all these spheres tolerates unreliability, then knowledge probably does too. For knowledge is a cognitive achievement, so absent special reason to think otherwise we should expect it to share the profile of achievements generally. But we've been given no special reason to suppose that knowledge differs in this regard. At this point, the balance of evidence has tipped against reliabilism, even if popular professional opinion has not.

However, some have suggested a possible alternative source of evidence for reliabilism. The suggestion taps into a deep and wide-ranging tradition in contemporary epistemology, which prizes a theory of knowledge that can explain our ordinary concept of knowledge and knowledge attributions (Austin 1956, Unger 1975, Dretske 1981, Stroud 1984, Vogel 1990, Craig 1990, DeRose 1995, Hawthorne 2004, Stanley 2005, Fantl & McGrath 2009, and Turri 2011).

The basic idea is that patterns in ordinary judgments and intuition support reliabilism. For example, John Greco claims that reliabilism is attractive because it “explains a wide range of our intuitions regarding what does and does not count as knowledge,” and we can “gain insight and understanding into what knowledge is” by “reflecting on our thinking and practices” surrounding knowledge (Greco 2010: 4-6). Fred Dretske defends an information-theoretic version of reliabilism partly on the grounds that it well fits “ordinary, intuitive judgments” (Dretske 1981: 92ff; see also p. 249, n. 8). Edward Craig advocates a version of reliabilism on the grounds that it “match-

es our everyday practice with the concept of knowledge as actually found,” that it is a “close fit to intuitive extension of ‘know’” (Craig 1990, pp. 3-4).<sup>2</sup> Ernest Sosa claims that reliabilism is a theoretical “correlate” of “commonsense” epistemology (Sosa 1991: 100). Goldman argues that one “proper task of epistemology is to elucidate our epistemic folkways,” and that an adequate epistemology must “have its roots in the concepts and practices of the folk” (Goldman 1993: 271). Goldman further argues that “folk” epistemology is an incipient form of reliabilism: “our inclinations” and “intuitions” about knowledge are keyed to reliability.

Call this *the proto-reliabilist hypothesis* about folk epistemology. This paper seeks evidence for the intriguing proto-reliabilist hypothesis. Taking a cue from Goldman, I treat the investigation as “a *scientific* task” and adopt the method of controlled experimentation typical in cognitive science (Goldman 1993: 272; see also Goldman & Olsson 2009: 35, who advocate explaining “commonsense epistemic valuations in a scientific fashion, especially by appeal to psychology”).

The proto-reliabilist hypothesis generates testable implications. There are at least four principal implications, either straightforwardly entailed by the hypothesis or suggested by reliabilists themselves. First and foremost, people should not count unreliably produced belief as knowledge. For example, if an ability produces only ten percent or thirty percent true beliefs, then the

---

2 Although Craig eschews the project of identifying necessary and sufficient conditions for the application of our ordinary knowledge concept, he also suggests that a good account of the concept will come close to fitting that format (Craig 1990: p. 5). For further discussion of the shortcomings of Craig’s intriguing view, see Turri in press a.

beliefs it produces should not be judged knowledge. Second, and closely related, vast and explicit differences in reliability should produce large differences in knowledge judgments. Third, reliabilists have claimed that their view captures the “fundamental intuition” that knowledge cannot be lucky (e.g. Goldman & Olsson 2009: 38; Heller 1999: 115; see also Heller 1995). Reliability is said to be at the heart of this “anti-luck intuition” (Becker 2006, Greco 2010: ch. 9; see also Pritchard 2012). So people should count an unreliably produced true belief as lucky and, therefore, not knowledge. Fourth, reliabilists, along with many others, have claimed that knowledge has certain counterfactual properties that are explained or captured by (very high) reliability (Goldman 2008; see also Dretske 1971, Nozick 1981, DeRose 1995, Sosa 1999, Vogel 2000). Reliability is said to be at the heart of these “counterfactual intuitions,” as we might call them. So people should attribute knowledge only if they attribute the counterfactual properties, and they should not count an unreliably produced true belief as having relevant counterfactual properties.

One prior study examined intuitions about a purported counterexample to reliabilism, Keith Lehrer’s Truetemp case (Lehrer 1990: 163-4; Swain, Alexander & Weinberg 2008). But those intuitions pertain to whether reliably produced true belief is *sufficient* for knowledge, whereas I’ll be concerned with whether it’s *necessary*. Aside from that one study, prior empirical work has not directly addressed the status of reliabilism in folk epistemology (for overviews, see Alexander 2012, Beebe 2012, Pinillos 2011, Beebe 2014).

The basic approach adopted here is to elicit judgments about simple cases that vary how reliably beliefs are formed. The main finding is that the proto-reliabilist hypothesis is clearly

false. Commonsense fully embraces the possibility of unreliable knowledge: rates of unreliable knowledge attribution sometimes reach 80-90%. Knowledge judgments are surprisingly insensitive to information about reliability: people attribute knowledge at similar rates whether the agent gets it right *ninety* percent of the time or *ten* percent of the time. Commonsense does share the intuition that knowledge isn't lucky, but the intuition is keyed to *ability*, reliable or unreliable, not reliability. Finally, reliabilists have badly mischaracterized the intuitive links among reliability, knowledge and various counterfactual judgments. I propose an alternative hypothesis, abilism, that better explains the findings and integrates them into a broader framework that also explains other surprising recent discoveries.

### **Preview of the experiments**

Experiments 1-6 address the first two principal implications of the proto-reliabilist hypothesis. They show that commonsense fully embraces the possibility of unreliable knowledge, and that people's knowledge judgments are unaffected by clear and explicit information about vast differences in reliability. Experiment 7 addresses the third principal implication of the proto-reliabilist hypothesis. It shows that reliability is not needed to rule out epistemic luck. Instead, what rules out luck is true belief produced by ability, regardless of whether it is reliable or unreliable. Experiments 8-9 address the fourth principal implication of the proto-reliabilist hypothesis. They show that reliabilists have simply misunderstood the intuitive connections among reliability, knowledge, and counterfactual judgments.

# Experiment 1

## Method

Participants (N = 113)<sup>3</sup> were randomly assigned to one of three groups, False/Ten/Ninety, in a between-subjects design. Each participant read a version of a single story about Alvin, who learns something important about the reliability or unreliability of his vision. Alvin comes to accept the results by reading them, thereby saliently basing his belief on vision. Here is the story (manipulation bracketed and separated by a slash):

[False/Ten/Ninety] Alvin wants to better understand how his own vision works. So he works with a team of cognitive scientists to develop a series of computerized tests that evaluate how well he visually detects things. After Alvin takes the tests, the computer prints out the results. Alvin reads the results: the central finding was that approximately [ten/ten/ninety] percent of Alvin's visual beliefs are true. This is an important fact about how Alvin's vision works, and he accepts the finding.

After reading the story, participants rated their agreement with the key test statement:

---

3 Aged 18-66 years, mean age = 32 years; 96% reporting English as a native language; 49 female. Participants were recruited and tested online (using Amazon Mechanical Turk and Qualtrics) and compensated \$0.35 for approximately 2 minutes of their time. Repeat participation was prevented. Eleven participants failed comprehension questions, but excluding them from the analysis does not affect the results reported below, so I retained their data to increase statistical power in this first study.



1. After reading the results, Alvin knows that approximately [seventy-five/ten/ninety] percent of his beliefs are true.

Participants in the False condition were presented with a knowledge statement that should be judged false because the purportedly known proposition is false and there is no textual basis for supposing that Alvin actually accepts it. Responses were collected on a standard 6-point Likert scale, anchored left-to-right with “Strongly disagree” (= 1), “Disagree,” “Slightly disagree,” “Slightly agree,” “Agree,” and “Strongly agree” (= 6). Participants were then taken to a separate screen to answer two comprehension questions from memory; the story did not remain at the top of the screen for these. Here are the questions (options in parentheses):

2. The test showed that approximately \_\_\_\_\_ percent of Alvin’s visual beliefs are true.  
[ten/seventy-five/ninety]
3. What happened in the story? [Alvin read the test results for himself./Someone read the test results to Alvin.]

Except for the elements of Likert scales, all options were rotated randomly. Participants never saw questions numbered or otherwise labeled.

I assume that the proto-reliabilist hypothesis predicts that people should somewhat disagree with the knowledge attribution in the Ten condition and somewhat agree in the Ninety condition. At the very least, if the proto-reliabilist hypothesis is true, then we should observe a mean difference of at least 1 on the six-point scale used — for example, the difference between “somewhat disagree” (= 3) and “somewhat agree” (= 4). *A priori* power analysis indicated that a sample size

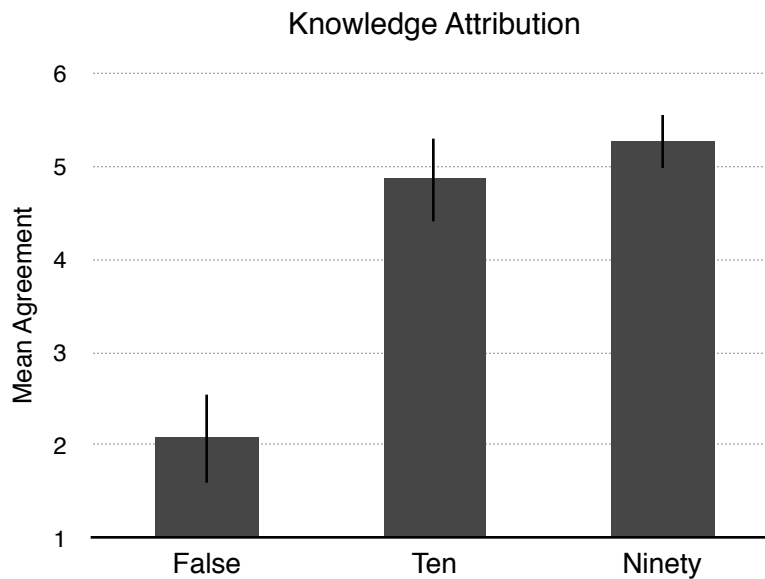
of 17 per group had sufficient power to detect a difference of this magnitude.<sup>4</sup> In this experiment, I more than *doubled* that sample size. So the present study had more than enough power to detect the expected difference, if it exists.<sup>5</sup>

## Results

Assignment to condition affected response to the knowledge statement.<sup>6</sup> But this was because

- 
- 4 Calculated using the software application G\*Power v. 3.1.5 (Faul, Erdfelder, Lang & Axel 2007). The calculation assumed a mean difference of 1, a common standard deviation of 1, and the conventional beta and alpha levels of .20 and .05.
  - 5 *Post hoc* power analysis revealed that the test's power to detect a mean difference of 1 was .97, which indicates that it had a 97% chance of detecting the difference, if it exists. The calculation used the observed common standard deviation of 1.11 and a sample size of 36 per group. *Post hoc* power analysis also revealed that the test was sufficiently powered to detect a mean difference of 0.66, assuming the conventional beta and alpha levels of .20 and .05.
  - 6 ANOVA,  $F(2, 110) = 67.93$ ,  $p < .001$ ,  $\eta_p^2 = .553$ . An initial omnibus ANOVA revealed no main or interaction effects for participant age: 2 equal groups, cut-point = 28 years;  $F(1, 79) = 0.68$ ,  $p = .412$ ) and no interaction effects for participant gender. However, it did indicate a main effect for participant gender, whereby women were more inclined to attribute knowledge than men,  $M = 4.27/3.77$ ,  $SD = 1.91/1.98$ ; gender,  $F(1, 110) = 5.48$ ,  $p = .021$ ,  $\eta_p^2 = .051$ . However, a follow-up independent samples t-test did not detect a gender difference,  $t(111) = -1.348$ ,  $p = .180$ ,  $MD = 0.50 [-1.23, 0.24]$ ; a nonparametric Mann-Whitney U test detected no difference either,  $U(113) = 1306$ ,  $Z = -1.57$ ,  $p = .117$ . Moreover, this gender-effect is not replicated in any other study reported here, which suggests that it might be a false positive (type I error): caution is warranted.

mean response in the False condition was significantly lower than in Ten and Ninety.<sup>7</sup> Mean response in Ten and Ninety didn't differ significantly.<sup>8</sup> The mode was “strongly disagree” (= 1) in False, “agree” (= 5) in Ten, and “strongly agree” (= 6) in Ninety.



**Fig 1.** Mean agreement with the knowledge statement on a 1 (strongly disagree) – 6 (strongly agree) scale. Error bars represent bootstrapped 95% confidence intervals.

## Discussion

These results provide initial evidence against the proto-reliabilist hypothesis. People seem perfectly happy to attribute unreliable knowledge. But we should not draw firm conclusions from a

---

7 False/Ten/Ninety:  $M = 2.07/4.86/5.27$ ,  $SD = 1.65/1.3/0.84$ . Independent samples t-test comparing False and Ten,  $t(74) = -8.05$ ,  $p < .001$ ,  $MD = -2.78$  [-3.47, -2.10]. Independent samples t-test comparing False and Ninety,  $t(60.7) = -10.615$ ,  $p < .001$ ,  $MD = -3.20$  [-3.78, -2.61].

8 Independent samples t-test,  $t(70) = -1.60$ ,  $p = .113$ ,  $MD = -0.41$  [-0.93, 0.10].

single experiment. Moreover, we might raise at least three concerns about this study.

First, Alvin is described in the context of a team of cognitive scientists. This might “rub off” on Alvin and make him sound more authoritative or reliable than intended. It also raises the possibility that although Alvin learns the tests results by reading them, people might just assume that the scientists are telling him the results too. This would provide a vision-independent basis for accepting the results. This should be ruled out.

Second, previous work on knowledge attributions shows that the mode of questioning can lead to surprising or counterintuitive results, which are not robust across different modes of questioning (Cullen 2010; Turri 2013). In particular, some researchers have suggested that the most appropriate way of probing for knowledge attributions involves a dichotomous choice between whether the person “knows” or “only believes/thinks” that a proposition is true (Starmans & Friedman 2012; Turri 2014; see also Weinberg, Nichols & Stich 2001). This probe seems to inhibit a tendency to treat “knowledge” as equivalent to “true belief,” which reliabilists deem irrelevant to their project (e.g. Goldman & Olsson 2009: sec. 1). Perhaps the present results are artifacts of using scaled responses. If so, then a dichotomous knowledge probe with an appropriate contrast will trigger people’s latent reliabilist tendencies. This should be tested.

Third, although the stories clearly stipulate the relevant percentages, participants did not rate Alvin’s reliability for themselves. It seems unlikely but it is still possible that the experimental manipulation was ineffective. Having participants rate Alvin’s reliability for themselves can address this. Relatedly, better evidence against the proto-reliabilist hypothesis would measure participants’ own assessment of reliability and check whether they attribute knowledge despite

actively categorizing Alvin as unreliable. This should be added to the experimental design.

The next experiment addresses all of these concerns.

## Experiment 2

### Method

New participants ( $N = 121$ )<sup>9</sup> were randomly assigned to one of four groups in a 2 (Percentage: Ten/Ninety)  $\times$  2 (Truth Value: False/True) between-subjects design, following procedures similar to those used above. The story was similar to the one used in Experiment 1, except that this time there were no cognitive scientists. The Percentage factor varied the explicitly stipulated percentage of Alvin's beliefs that are true (ten versus ninety). The Truth Value factor varied whether the test results are false or true. Here is the story:

Alvin wants to better understand how his own vision works. So he undergoes a series of computerized tests that evaluate how well he visually detects things. After the tests are complete, Alvin reads the results, which indicate that approximately [ten/ninety] percent of his visual beliefs are true. Alvin judges these to be important results about how his vision works. [However/Moreover], the results are completely [in/valid] and [in/accurate].

---

9 Aged 18-72 years, mean age = 33 years; 94% reporting English as a native language; 42 female. I excluded data from nineteen participants who failed comprehension questions, but including them does not affect the results reported below.

Participants then responded to an open knowledge probe:

1. After reading the results, Alvin \_\_\_\_\_ that approximately [ten/ninety] percent of his visual beliefs are true. (knows/only thinks)<sup>10</sup>

Participants then went to a new screen where they rated their agreement, on the same 6-point Likert scale used in Experiment 1, with the following statement:

2. Most of Alvin's visual beliefs are true.

Finally, participants answered from memory the same two comprehension questions as in Experiment 1.

## **Results**

If the proto-realist hypothesis is true and the concerns raised about the earlier results are legitimate, then we should observe two things. First, people should attribute knowledge significantly more in the True Ninety condition than in the True Ten condition. Second, people should attribute knowledge at rates below chance in the True Ten condition. Aside from either of those predictions, we should expect people in each False condition to attribute knowledge at rates below chance.

---

<sup>10</sup> Participants also rated their confidence in their answer to the dichotomous knowledge probe. Combining the dichotomous response with reported confidence resulted in the same pattern of results as for the dichotomous response alone, so I omit discussion of it.

None of the predictions turned out true.<sup>11</sup> First, in the True Ninety condition, 87% of participants attributed knowledge, which was not significantly more than the 75% who attributed knowledge in the True Ten condition.<sup>12</sup> Second, participants in the True Ten condition attributed knowledge at rates significantly *above* chance.<sup>13</sup> Third, only 13% of participants in the False Ten condition attributed knowledge, which is far below chance,<sup>14</sup> but a surprisingly high 39% attributed knowledge in the False Ninety condition, which does not differ significantly from chance.<sup>15</sup> (Note: this puzzling finding in False Ninety does not replicate in Experiment 3, so it might be an anomaly.)<sup>16</sup> Knowledge attribution in False Ninety was significantly higher than in False Ten.<sup>17</sup>

---

11 Preliminary analyses revealed no effect of participant age or gender on knowledge judgments.

12 Fisher's exact test,  $p = .318$ .

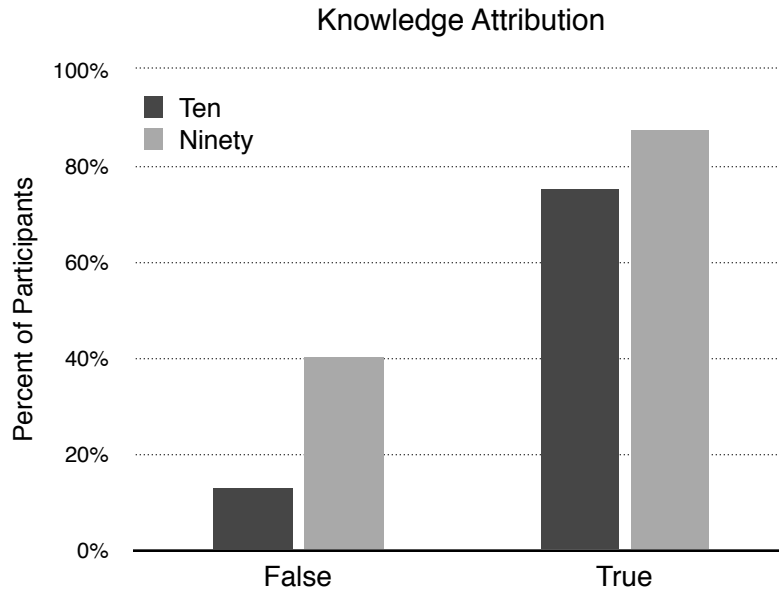
13 Chi-square goodness of fit test,  $\chi^2(1, 28) = 7.00$ ,  $p = .008$ .

14 Chi-square goodness of fit test,  $\chi^2(1, 30) = 16.13$ ,  $p < .001$ .

15 Chi-square goodness of fit test,  $\chi^2(1, 31) = 1.58$ ,  $p = .209$ .

16 However, also compare the finding in Experiment 6, where increased reliability made it more likely that people attribute knowledge to someone who "just guesses" the correct answer.

17 Fisher's exact test,  $p = .04$ , Cramer's  $V = .288$ .



**Fig. 2.** Experiment 2: Percentage of participants attributing knowledge.

Question 2 serves as a manipulation check. Response to this question cannot be interpreted very well in the False conditions because participants were given no positive information about Alvin’s reliability, only negative information about what the relevant percentage is not. But this is not worrisome because it is beside the point of including the manipulation check. The check was included to measure how participants in True conditions assess Alvin’s reliability. More specifically, it was included to answer two questions. First, do participants rate Alvin’s reliability significantly higher in True Ninety than in True Ten? Second, do participants in True Ten attribute knowledge despite rating Alvin as unreliable?

The answer to both questions is “yes.” On the one hand, participants rated Alvin’s reliability



significantly higher in True Ninety than in True Ten.<sup>18</sup> Mean reliability rating was significantly below the theoretical midpoint (= 3.5) in True Ten, and it was significantly above the midpoint in True Ninety.<sup>19</sup> On the other hand, among the twenty-one people in True Ten who attributed knowledge, mean reliability rating was significantly below the midpoint.<sup>20</sup> The modal response in this group was “strongly disagree” (= 1). Moreover, mean reliability rating in this group did not differ significantly from the mean for those in True Ten who denied knowledge.<sup>21</sup>

## Discussion

These results replicate the main findings from Experiment 1. Once again, people were happy to attribute unreliable knowledge, and explicitly stipulating radically different levels of reliability had no effect on knowledge attribution. Moreover, I obtained these results using a different dependent measure; I confirmed that the reliability manipulation was effective; I observed that people attribute knowledge while actively categorizing the knower as unreliable; and I found that participants in a matched control condition overwhelmingly deny knowledge. One potential concern about this study is that participants might have distinguished between *visually detecting*

---

18 True Ten/True Ninety:  $M = 2.54/5.25$ ,  $SD = 1.69/0.57$ . Independent samples t-test,  $t(32.3) = -8.12$ ,  $p < .001$ ,  $MD = -2.71 [-3.39, -2.03]$ .

19 One sample t-tests. True Ten:  $t(27) = -3.02$ ,  $p = .005$ ,  $MD = -0.96 [-1.62, -0.31]$ . True Ninety:  $t(31) = 17.43$ ,  $p < .001$ ,  $MD = 1.75 [1.55, 1.95]$ .

20  $M = 2.48$ ,  $SD = 1.63$ . One sample t-test,  $t(20) = -2.88$ ,  $p = .009$ ,  $MD = -1.02 [-1.77, -0.28]$ .

21 Attribute/Deny:  $M = 2.48/2.71$ ,  $SD = 1.63/1.98$ . Independent samples t-test,  $t(26) = 0.32$ ,  $p = .753$ ,  $MD = 0.24 [-1.30, 1.78]$ .

*things* and *reading*. If they did, then they might have agreed that Alvin is unreliable at visually detecting things but not reading. The next experiment addresses this concern.

### Experiment 3

#### Method

New participants (N = 123)<sup>22</sup> were randomly assigned to one of four groups in a 2 (Percentage: Ten/Ninety) × 2 (Truth Value: False/True) between-subjects design, following procedures similar to those used above. The story was similar to the one used in Experiment 2, except that this time the test pertained specifically to reading skills and beliefs formed through reading (“reading beliefs”). Here is the story:

Alvin wants to better understand his reading skills. So he undergoes a series of computerized tests that evaluate how well he reads. After the tests are complete, Alvin reads the results, which indicate that approximately [ten/ninety] percent of his beliefs formed through reading (“reading beliefs”) are true. Alvin judges these to be important results about his reading skills. [However/Moreover], the results are completely [in/valid] and [in/accurate].

Participants then responded to an open knowledge probe:

---

<sup>22</sup> Aged 19-67 years, mean age = 34 years; 94% reporting English as a native language; 55 female.

1. Alvin \_\_\_\_\_ that approximately ten percent of his reading beliefs are true.. (knows/  
only thinks)

Participants then went to a new screen where they rated their agreement, on the same 6-point scale used earlier, with the following statement:

2. Most of Alvin's reading beliefs are true.

## Results

If the concern raised about Experiment 2 was well founded, then the results this time will be more favorable to the proto-reliabilist hypothesis. By contrast, if the results are broadly similar to those observed in Experiment 2, then the earlier concern was not well founded.

The results replicated the findings from Experiment 2.<sup>23</sup> In the True Ninety condition, 74% of participants attributed knowledge, which was exactly the same as in the True Ten condition.<sup>24</sup> Participants in the True Ten condition attributed knowledge at rates significantly above chance.<sup>25</sup> Only 17% of participants in the False Ten condition attributed knowledge, which is far below chance<sup>26</sup> and not significantly different from the 23% who attributed knowledge in the False Ninety condition.<sup>27</sup> (Note: this fails to replicate the puzzling finding in False Ninety in Experi-

---

23 Preliminary analyses revealed no effect of participant age or gender on knowledge judgments.

24 Fisher's exact test,  $p = 1$ .

25 Chi-square goodness of fit test,  $\chi^2(1, 31) = 7.26, p = .007$ .

26 Chi-square goodness of fit test,  $\chi^2(1, 30) = 13.33, p < .001$ .

27 Fisher's exact test,  $p = .561$ .

ment 2.) Participants rated Alvin's reliability significantly higher in True Ninety than in True Ten.<sup>28</sup> Mean reliability rating was trending below the theoretical midpoint (= 3.5) in True Ten, and it was significantly above the midpoint in True Ninety.<sup>29</sup> On the other hand, among the twenty-three people in True Ten who attributed knowledge, mean reliability rating was significantly below the midpoint.<sup>30</sup> The modal response in this group was "disagree" (= 2). Moreover, mean reliability rating in this group was non-significantly higher than the mean for those in True Ten who denied knowledge.<sup>31</sup>

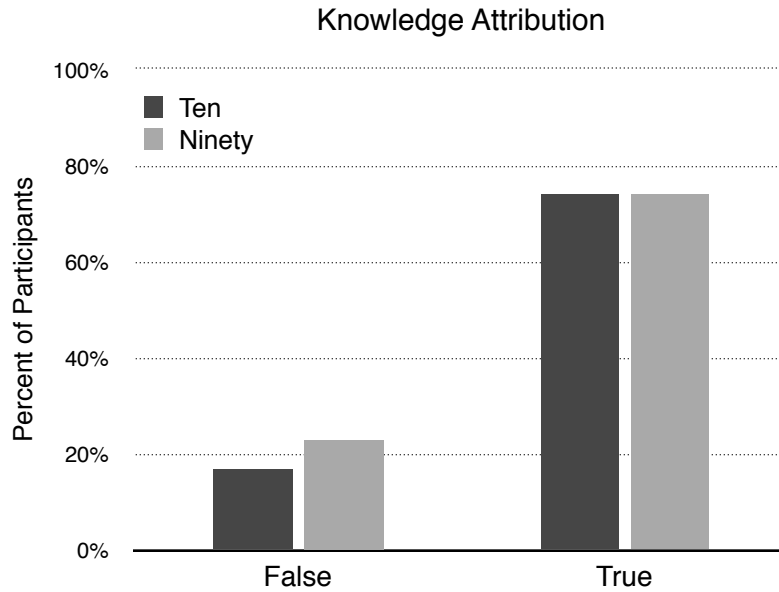
---

28 True Ten/True Ninety:  $M = 2.94/5.26$ ,  $SD = 1.73/0.73$ . Independent samples t-test,  $t(40.32) = -6.89$ ,  $p < .001$ ,  $MD = -2.32 [-3.00, -1.64]$ .

29 One sample t-tests. True Ten:  $t(30) = -1.82$ ,  $p = .079$ ,  $MD = -0.57 [-1.20, 0.07]$ . True Ninety:  $t(30) = 13.43$ ,  $p < .001$ ,  $MD = 1.76 [1.49, 2.03]$ .

30  $M = 2.65$ ,  $SD = 1.75$ . One sample t-test,  $t(22) = -2.33$ ,  $p = .030$ ,  $MD = -0.85 [-1.60, -0.09]$ .

31 Attribute/Deny:  $M = 2.65/3.75$ ,  $SD = 1.75/1.49$ . Independent samples t-test,  $t(29) = 1.58$ ,  $p = .124$ ,  $MD = -1.10 [-0.32, 2.52]$ .



**Fig. 3.** Experiment 3: Percentage of participants attributing knowledge.

## Discussion

These results replicate the main findings from Experiment 2. The next experiment investigates whether these findings generalize to other narrative contexts (e.g. no computerized tests) and cognitive faculties (e.g. memory).

## Experiment 4

### Method

New participants ( $N = 138$ )<sup>32</sup> were randomly assigned to one of four groups in a 2 (Quality: Un-

<sup>32</sup> Aged 18-72 years, mean age = 33 years; 94% reporting English as a native language; 42 female. I excluded data from nineteen participants who failed comprehension questions, but including them does not affect the results reported below.

reliable/Reliable) × 2 (Truth Value: False/True) between-subjects design, following procedures similar to those used above. The story was completely new and focused on Alvin's memory rather than his vision. The Quality factor varied whether Alvin's memory is qualitative described as "very unreliable" or "very reliable." The Truth Value factor varied the truth-value of the relevant claim. Here is the story:

Alvin has been out running errands all afternoon. He has a very [un/reliable] memory. This morning his wife told him that he should [not] stop at the dry cleaners, because their dry cleaning [is ready to be picked up/will not be ready for a couple days]. On his way home, Alvin stops at the dry cleaners.

Participants then responded to an open knowledge probe:

1. Alvin \_\_\_\_\_ that he should stop at the dry cleaners. (knows/only thinks)

Participants then went to a new screen where they rated their agreement, on the same 6-point Likert scale from earlier experiments, with the following statement:

2. Alvin's memory is reliable.

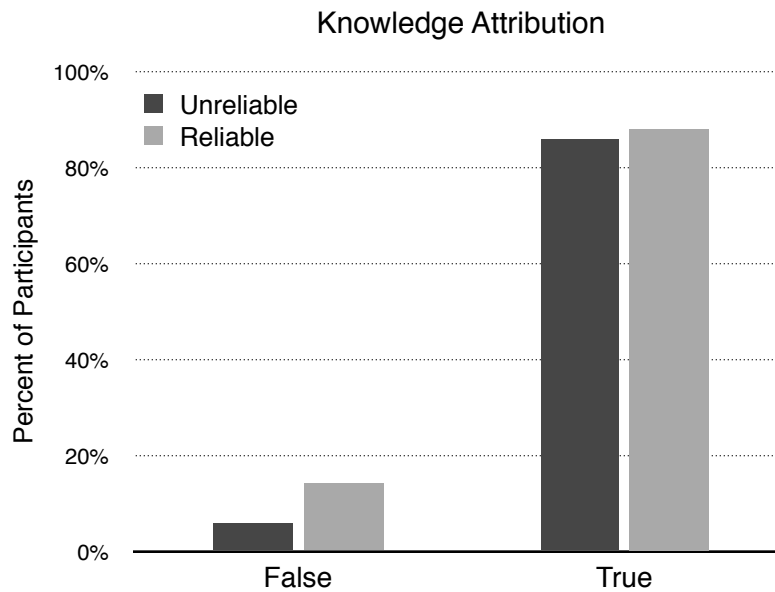
Finally, participants answered from memory a comprehension question:

3. Alvin was out all afternoon \_\_\_\_\_. [running errands/delivering dry cleaning]

## **Results**

The results were very similar to Experiments 2-3. First, in the Reliable True condition, 88% of participants attributed knowledge, which was not significantly more than the 86% who attributed

knowledge in the Unreliable True condition.<sup>33</sup> Second, participants in the Unreliable True condition attributed knowledge at rates significantly above chance.<sup>34</sup> Third, only 6% of participants in the Unreliable False condition attributed knowledge, which was not significantly different from the 14% who attributed knowledge in the Reliable False condition.<sup>35</sup> In both False conditions, knowledge attribution was far below chance.<sup>36</sup>



**Fig. 4.** Experiment 4: Percentage of participants attributing knowledge.

Question 2 again served as a manipulation check, included to answer two questions. First,

---

33 Fisher's exact test,  $p = 1$ .

34 Chi-square goodness of fit test,  $\chi^2(1, 36) = 18.78, p < .001$ .

35 Fisher's exact test,  $p = .429$ .

36 Chi-square goodness of fit tests. Unreliable False:  $\chi^2(1, 33) = 25.49, p < .001$ . Reliable False:  $\chi^2(1, 35) = 17.86, p < .001$ .

do participants rate Alvin's reliability significantly higher in Reliable conditions than in Unreliable conditions? Second, do participants in the Unreliable True condition attribute knowledge despite rating Alvin as unreliable? The answer to both questions is "yes." On the one hand, participants rated Alvin's reliability significantly higher in Reliable conditions than in Unreliable conditions.<sup>37</sup> On the other hand, among the 31 people in Unreliable True who attributed knowledge, mean reliability rating was significantly below the midpoint (= 3.5).<sup>38</sup> The modal response in this group was "disagree" (= 2). Moreover, mean reliability rating in this group did not differ significantly from the mean rating of those in Unreliable True who denied knowledge.<sup>39</sup>

## Discussion

These results replicate and generalize the findings from the earlier experiments. People overwhelmingly attributed unreliable knowledge. People attributed knowledge while actively categorizing the knower as unreliable. Participants in a very closely matched control condition overwhelmingly denied knowledge. Explicitly stipulating radically different levels of reliability had no effect on knowledge attribution. The reliability manipulation was effective. I observed all

---

37 Unreliable False/Reliable False:  $M = 1.73/2.94$ ,  $SD = 1.26/1.41$ ; independent samples t-test,  $t(66) = -3.74$ ,  $p < .001$ ,  $MD = -1.22 [-1.87, -0.57]$ . Unreliable True/Reliable True:  $M = 2.17/5.56$ ,  $SD = 1.3/0.71$ ; independent samples t-test,  $t(54.6) = -13.69$ ,  $p < .001$ ,  $MD = -3.39 [-3.89, -2.90]$ .

38 One sample t-test,  $t(30) = -4.94$ ,  $p < .001$ ,  $MD = -1.77 [-1.66, -0.69]$ .

39 Knows/only thinks:  $M = 2.32/1.20$ ,  $SD = 1.3/0.45$ . Independent samples t-test,  $t(34) = 1.86$ ,  $p = .072$ ,  $MD = 1.12 [-0.11, 2.35]$ .



these things while using a very different cover story that focused on a different cognitive faculty (memory). One potential concern about this experiment is that participants in True conditions might have misunderstood the knowledge attribution to pertain to an obligation that Alvin has to his wife, in light of their earlier conversation.<sup>40</sup> The next experiment addresses this possibility.

## Experiment 5

### Method

New participants (N = 120)<sup>41</sup> were randomly assigned to one of four groups in a 2 (Quality: Unreliable/Reliable) × 2 (Truth Value: False/True) between-subjects design, following procedures similar to those used above. The story was similar to the one used in Experiment 4, except this time Alvin's errand pertained to a prescription he needed to pick up, rather than his wife telling him to do something. Here is the story:

Alvin is very [unreliable/reliable] at remembering driving directions. Today he is visiting a friend in an unfamiliar town. Alvin needs to pick up a prescription while he is there, so his friend gives him directions to the pharmacy. On the way, Alvin needs to make a [left/right] turn at an intersection. Alvin gets to the intersection

---

<sup>40</sup> An anonymous referee raised this alternative reading.

<sup>41</sup> Aged 18-72 years, mean age = 33 years; 94% reporting English as a native language; 42 female. I excluded data from nineteen participants who failed comprehension questions, but including them does not affect the results reported below.

and turns right.

Participants then responded to an open knowledge probe, which can only be interpreted as pertaining to the correct driving directions:

1. When he got to the intersection, Alvin \_\_\_\_\_ that he should turn right to get to the pharmacy. (knew/only thought)

Participants then went to a new screen where they rated their agreement, on the same 6-point Likert scale from earlier experiments, with the following statement:

2. When it comes to driving directions, Alvin's memory is reliable.

## Results

The results were very similar to Experiment 4. First, in the Reliable True condition, 80% of participants attributed knowledge, which was not significantly more than the 77% who attributed knowledge in the Unreliable True condition.<sup>42</sup> Second, participants in the Unreliable True condition attributed knowledge at rates significantly above chance.<sup>43</sup> Third, only 13% of participants in the Unreliable False condition attributed knowledge, which was not significantly different from the 7% who attributed knowledge in the Reliable False condition.<sup>44</sup> In both False condi-

---

42 Fisher's exact test,  $p = 1$ .

43 Chi-square goodness of fit test,  $\chi^2(1, 30) = 8.53$ ,  $p = .003$ .

44 Fisher's exact test,  $p = .671$ .

tions, knowledge attribution was far below chance.<sup>45</sup> Participants rated Alvin’s reliability significantly higher in True Reliable than in True Unreliable.<sup>46</sup> Mean reliability rating was non-significantly below the theoretical midpoint (= 3.5) in True Unreliable, and it was significantly above the midpoint in True Reliable.<sup>47</sup>



**Fig. 5.** Experiment 5: Percentage of participants attributing knowledge.

## Discussion

These results replicate the main findings from Experiment 4. The next experiment generalizes

45 Chi-square goodness of fit tests. Unreliable False:  $\chi^2(1, 30) = 16.13, p < .001$ . Reliable False:  $\chi^2(1, 30) = 22.53, p < .001$ .

46 True Unreliable/True Reliable:  $M = 3.03/5.43, SD = 1.61/0.63$ . Independent samples t-test,  $t(37.60) = -7.62, MD = 2.40 [-3.04, -1.76]$ .

47 One sample t-tests. True Unreliable:  $t(29) = -1.59, p = .123, MD = -0.47 [-1.07, 0.13]$ . True Reliable:  $t(29) = 16.91, p < .001, MD = 1.93 [1.70, 2.17]$ .

these results even further.

## Experiment 6

### Method

New participants ( $N = 133$ )<sup>48</sup> were randomly assigned to one of four groups in a 2 (Source: Guess/Ability)  $\times$  2 (Percent: Thirty/Ninety) between-subjects design, following procedures similar to those used above. The story was completely new and focused on Carolyn, who has a rare ability to read words flashed on a screen for only 120 milliseconds. Most people correctly report less than one percent of such words, but Carolyn's ability allows her to perform better. The Percent factor varied whether Carolyn's reports were correct thirty or ninety percent of the time. On the present occasion, when the word "Corn" is flashed on the screen, Carolyn reports, "Corn." The Source factor varied whether Carolyn was just guessing or exercising her ability this time when she got it right. Here is the the story:

Cognitive scientists study the limits of human perception. One experiment tested whether humans could read words flashed on a screen for 120 milliseconds. This is so fast that almost all participants are reduced to guessing. Less than one per-

---

48 Aged 19-69 years, mean age = 32 years; 97% reporting English as a native language; 54 female. Thirteen participants didn't correctly answer the manipulation check correctly. But, as with every other experiment reported in this paper, excluding them from the analysis didn't affect the pattern of results reported in the main text, so I retained their responses to increase statistical power.

cent of their reports are correct. ¶<sup>49</sup> But Carolyn doesn't have to guess. She has a special ability to detect words flashed that fast. Scientists think it has something to do with a unique feature of her optical nerves. Astonishingly, [thirty/ninety] percent of Carolyn's reports are correct. ¶ Today Carolyn took the flash test again. The word "corn" flashed on the screen and Carolyn reported, "Corn." She was [exercising her special ability/just guessing this time] when she got it right.

After reading the story, participants responded to an open knowledge probe.

1. Carolyn \_\_\_\_\_ that the word was "Corn". (knows/only thinks)

Participants then rated their confidence in their answer to the knowledge probe on a 1 ("not at all confident") – 6 ("completely confident") scale. Participants then went to a new screen and answered from memory a manipulation check:

2. Scientists have found that when Carolyn is taking the flash-word test, approximately \_\_\_\_\_ of her reports are correct. (one/thirty/ninety percent).

## Results

The Percent manipulation was extremely effective.<sup>50</sup> To analyze the results for knowledge judgments, I combined people's response to the dichotomous knowledge probe with their confidence. For each participant, an answer of 'knows' was scored +1 and 'only thinks' was scored -1. This

---

49 Indicates paragraph break on the participant's screen.

50 Ninety percent of participants correctly identified Carolyn's level of reliability, chi-square goodness of fit test,  $\chi^2(1, 133) = 79.77, p < .001$ .

score was then multiplied by the participant's reported confidence (1-6) to generate a *knowledge score* on a scale that ran -6 (maximum knowledge denial) to +6 (maximum knowledge attribution) (following the procedures of Starmans & Friedman 2012).

Thirty percent accuracy is far less than reliabilism requires for knowledge. So if the proto-reliabilist hypothesis is true, then knowledge scores should be low in the Thirty conditions. Moreover, the difference between thirty and ninety percent accuracy is unmistakably large, so if the proto-reliabilist hypothesis is true, then we should observe a large difference between Thirty Ability and Ninety Ability.

Neither prediction was true. Source significantly affected knowledge scores,<sup>51</sup> but Percent did not.<sup>52</sup> Knowledge scores were significantly above the neutral midpoint (= 0) in both Thirty Ability and Ninety Ability, whereas they were significantly below the midpoint in Thirty Guess and Ninety Guess.<sup>53</sup> Knowledge scores did not differ between Thirty Guess and Ninety Guess,<sup>54</sup> or between Thirty Ability and Ninety Ability.<sup>55</sup>

---

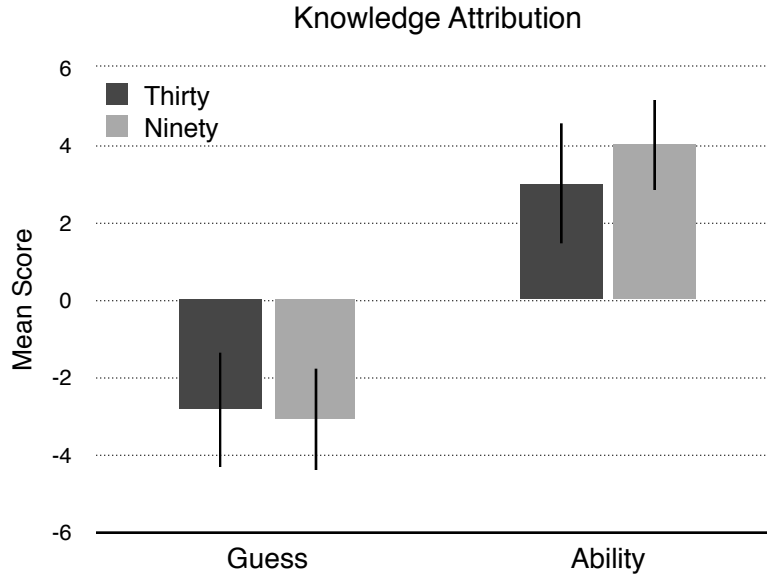
51 ANOVA,  $F(1, 129) = 91.01$ ,  $p < .001$ ,  $\eta_p^2 = .414$ .

52 ANOVA,  $F(1, 129) = 0.30$ ,  $p = .535$ ,  $\eta_p^2 = .003$ .

53 One sample t-tests: Thirty Ability,  $M = 3.00$ ,  $SD = 4.28$ ,  $t(31) = 3.97$ ,  $p < .001$ ,  $MD = 3.00$  [1.46, 4.54]. Ninety Ability,  $M = 4.09$ ,  $SD = 3.32$ ,  $t(33) = 7.19$ ,  $p < .001$ ,  $MD = 4.09$  [2.93, 5.24]. Thirty Guess,  $M = -2.85$ ,  $SD = 4.37$ ,  $t(32) = -3.74$ ,  $p < .001$ ,  $MD = -2.85$  [-4.40, -1.30]. Ninety Guess,  $M = -3.09$ ,  $SD = 3.70$ ,  $t(33) = -4.86$ ,  $p < .001$ ,  $MD = -3.09$  [-4.38, -1.80].

54 Independent samples t-test,  $t(65) = 0.24$ ,  $p = .809$ ,  $MD = 0.24$  [-1.74, 2.21].

55 Independent samples t-test,  $t(58.4) = -1.15$ ,  $p = .255$ ,  $MD = -1.09$  [-2.98, 0.81].



**Fig. 6.** Experiment 6: mean knowledge scores, derived for each participant by multiplying their dichotomous answer (scored +1 for ‘knows’, -1 for ‘only thinks’) by their reported confidence (1 = “not at all confident” to 6 = “completely confident”). Scales ran from -6 to +6. Error bars represent bootstrapped 95% confidence intervals.

## Discussion

These results once again replicate the main findings that people readily attribute unreliable knowledge, and that explicitly stipulating radically different levels of reliability doesn’t significantly affect knowledge attribution. These results are exactly the opposite of what we would expect if the proto-reliabilist hypothesis was true.

One result from the present study merits special emphasis: the difference between knowledge judgments in Thirty Guess and Thirty Ability, which is predicted by abilism but not reliabilism. In each condition the agent has an unreliable ability. In Thirty Ability she gets the truth because she exercised her ability, whereas in Thirty Guess she gets the truth by just guessing. In

other words, the two cases don't even differ in whether the agent possesses the unreliable ability, only in whether she exercises it. Yet the accompanying difference in knowledge attributions is not only statistically significant but also extremely large (approximately 80% compared to 25%).

By this point, it is clear that the first two principal implications of the proto-reliabilist hypothesis are false. The first principal implication is that people will not attribute unreliable knowledge. The second implication is that vast and explicit differences in reliability will produce large differences in knowledge judgments. In the next experiment, I turn my attention to the third principal implication, namely, that people will count an unreliably produced true belief as lucky and, therefore, not knowledge.

## **Experiment 7**

### **Method**

New participants ( $N = 226$ )<sup>56</sup> were randomly assigned to one of three groups, Control/Ability/Guess, in a between-subjects design, following procedures similar to those used above. The same basic story from the Thirty conditions in Experiment 6 was used. On the present occasion, when the word 'corn' is flashed on the screen, Carolyn reports, "Corn." In the story for the Ability

---

56 Aged 18-72 years, mean age = 32 years; 97% reporting English as a native language; 89 female. I recruited a higher number of participants per condition this time mainly because I had no way of estimating the potential difference between Control and Ability, which I was interested in. So I doubled the participants per condition. As it turns out, the additional statistical power was inadequate to detect a significant difference between Control and Ability.



condition, Carolyn was exercising her ability when she got it right (exactly the same as Thirty Ability from Experiment 6); in the story for the Guess condition, she was just guessing (exactly the same as Thirty Guess from Experiment 6); and in the story for the Control condition, neither her ability nor guessing is identified (the final sentence: “The person conducting the test noted that she got it right”). Participants then responded to the same open knowledge probe as in Experiment 6 and rated their confidence on the same 1–6 scale. Finally, participants then went to a new screen where they responded to an open explanation probe:

4. Carolyn got the right answer because of \_\_\_\_\_. (her ability/good luck)

Participants then rated their confidence in their answer to the explanation probe on the same 1–6 scale.

## **Results**

Thirty percent accuracy is far less than reliabilism requires for knowledge. So if the proto-realist hypothesis is true, then knowledge scores should be low in all three conditions. Moreover, if reliabilists are right about the “fundamental” anti-luck intuition, then explanation scores should be low too (i.e. participants should attribute the truth of Carolyn’s belief to luck). By contrast, if abilism better captures the ordinary concept of knowledge and we can gain knowledge through unreliable abilities, then knowledge scores will be high in the Ability condition, and explanation scores should be high too (i.e. participants should attribute the truth of Carolyn’s belief to ability). Moreover, if abilism is true, then merely having the unreliable ability alongside the true belief will *not* be enough for Carolyn to gain knowledge: the ability also has to *produce* the true

belief. Thus abilism predicts a large difference between knowledge scores in the Ability and Guess conditions.

To analyze the results, I calculated a knowledge score for each participant exactly the same way as in Experiment 6. Analogously, each participant's response to the explanation probe was scored either +1 ('her ability') or -1 ('good luck') and multiplied by the confidence rating, yielding an *explanation score* on the same -6 to +6 scale. Using these weighted scores allows for more powerful statistical tests and procedures.

Assignment to condition significantly affected knowledge scores.<sup>57</sup> But this was because knowledge scores in Guess were significantly lower than scores in Control and Ability.<sup>58</sup> Knowledge scores in Control and Ability didn't differ.<sup>59</sup> Knowledge scores were significantly above the neutral midpoint (= 0) in Control and Ability, whereas they were significantly below the mid-

---

57 ANOVA,  $F(2, 223) = 50.95$ ,  $p < .001$ ,  $\eta_p^2 = .314$ . An initial omnibus ANOVA revealed no main or interaction effects for participant age or gender: gender,  $F(1, 214) = 0.54$ ,  $p = .464$ ; age,  $F(1, 214) = 0.09$ ,  $p = .771$ .

58 Control/Ability/Guess:  $M = 2.21/2.19/-3.58$ ,  $SD = 4.20/4.37/3.60$ . Independent samples t-test comparing Control and Guess,  $t(145.14) = 9.10$ ,  $p < .001$ ,  $MD = 5.79$  [4.54, 7.05]; comparing Ability and Guess,  $t(143.12) = 8.84$ ,  $p < .001$ ,  $MD = 5.76$  [4.48, 7.05]. The effect size for the Ability/Guess comparison is extremely large: the eta squared statistic is .353, which indicates that over 35% of variance in knowledge judgments is explained by the difference between getting the truth through ability rather than luck.

59 Independent samples t-test,  $t(148) = 0.04$ ,  $p = .970$ .

point in Guess.<sup>60</sup> Assignment to condition also significantly affected explanation scores.<sup>61</sup> Explanation scores in Control and Ability were significantly above the midpoint, whereas they were significantly below the midpoint in Guess.<sup>62</sup>

Knowledge scores and explanation scores were very strongly positively correlated<sup>63</sup> and, moreover, explanation scores mediated the effect of condition on knowledge scores, suggesting that people attributed knowledge *because* they attributed Carolyn's true belief to her unreliable

---

60 One sample t-tests: Control,  $t(74) = 4.57$ ,  $p < .001$ , MD = 2.21 [1.25, 3.18]. Ability,  $t(74) = 4.37$ ,  $p < .001$ , MD = 2.19 [1.18, 3.19]. Guess,  $t(75) = -8.66$ ,  $p < .001$ , MD = -3.58 [-4.40, -2.76].

61 ANOVA,  $F(2, 223) = 77.51$ ,  $p < .001$ ,  $\eta_p^2 = .41$ . An initial omnibus ANOVA revealed no main or interaction effects for participant age (two equal groups, cut-point = 29 years) or gender: gender,  $F(1, 214) = 0.09$ ,  $p = .770$ ; age,  $F(1, 214) = 0.58$ ,  $p = .447$ .

62 Control/Ability/Guess:  $M = 4.61/4.28/-1.38$ ,  $SD = 2.00/2.62/4.71$ . One sample t-tests: Control,  $t(74) = 19.92$ ,  $p < .001$ , MD = 4.61 [4.15, 5.07]. Ability,  $t(74) = 14.16$ ,  $p < .001$ , MD = 4.28 [3.68, 4.88]. Guess,  $t(75) = -2.56$ ,  $p = .013$ , MD = -1.38 [-2.46, -0.31].

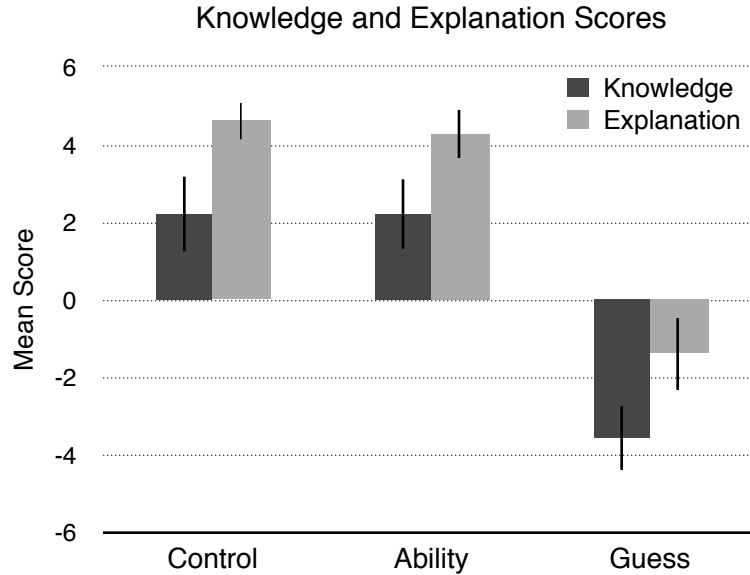
63 Pearson's correlation,  $r = .622$ ,  $N = 226$ ,  $p < .001$ .

ability.<sup>64</sup> Focusing on just the dichotomous probes, attributing Carolyn's true belief to ability rather than luck increased the odds of attributing knowledge by over 3800% (or a factor of 38.33).<sup>65</sup>

---

64 As determined by a bootstrap mediation analysis (Hayes 2013). Because knowledge scores and explanations scores were basically the same in the Control and Ability conditions, I first collapsed these into a single group (Ability\*) for purposes of analysis. I then entered condition as the independent variable (dummy coded, Guess = 0, Ability\* = 1), knowledge score as the dependent variable, and explanation score as the potential mediator: 95% CI for the indirect effect = 2.11 to 4.12; ratio of indirect to direct effect = 1.04, 95% CI = 0.53 to 2.94. The reverse mediation model — with condition as the independent variable, explanation score as the dependent variable, and knowledge score as the potential mediator — was also significant: 95% CI for the indirect effect = 1.38 to 2.73; ratio of indirect to direct effect = 0.51, 95% CI = 0.28 to 0.96. But I assume that the causal model presupposed by this latter mediation analysis is implausible. I will not pursue the matter further here.

65 Binary logistic regression, Beta = 3.645, SE = 0.62, Wald = 35.10, df = 1, p < .001, odds ratio = 38.33 [11.47, 128.13]. The model contained answer to the dichotomous explanation probe as the predictor and answer to the dichotomous knowledge probe as the outcome. The model was statistically significant,  $\chi^2(1, N = 226) = 77.59$ , explained between 29.1% and 38.8% of the variance in knowledge attribution, and correctly classified 75% of cases.



**Fig. 7.** Experiment 7: Mean knowledge scores and explanation scores (within-subjects). Knowledge scores were derived for each participant by multiplying their dichotomous answer (scored +1 for ‘knows’, -1 for ‘only thinks’) by their reported confidence (1 = “not at all confident” to 6 = “completely confident”). Explanation scores were derived for each participant by multiplying their dichotomous answer (scored +1 for ‘ability’, -1 for ‘luck’) by their reported confidence. Scales ran from -6 to +6. Error bars represent bootstrapped 95% confidence intervals.

## Discussion

These results show that reliabilists have mischaracterized the “fundamental” intuitive relationship between knowledge and luck. Most people do not share the intuition that reliability specifically is what rules out problematic luck. Instead, it seems that reaching the truth through ability, reliable or not, suffices to rule out luck, which falsifies the third principal implication of the proto-reliabilist hypothesis. Once again, the results also replicate and further generalize the main finding from earlier experiments, namely, that people readily attribute unreliable knowledge.

Overall, the results were exactly as we would expect if abilism adequately captured the ordinary concept of knowledge.

The results also replicate an important finding from Experiment 6: the difference between knowledge judgments in the Ability and Guess conditions, which is predicted by abilism but not reliabilism. In each condition the agent has an unreliable ability. In Ability she gets the truth because she exercised her ability, whereas in Guess she gets the truth by just guessing. In other words, the two cases don't even differ in whether the agent possesses the unreliable ability, only in whether she exercises it. Yet the accompanying difference in knowledge attributions was not only statistically significant but also very large (approximately 70% compared to 15%).

A potentially important null result is the striking similarity between scores in the Control and Ability conditions. This suggests a general lesson about commonsense cognitive evaluations: when someone forms a true belief and has a cognitive ability relevant to detecting the truth, the default assumption is that she formed the true belief through ability and, as a result, knows.

The final two experiments test the fourth and final principal implication of the proto-reliabilist hypothesis, regarding intuitive connections among reliability, knowledge, and counterfactuals.

## **Experiment 8**

Many philosophers, reliabilists prominently among them, find it intuitive that knowledge has certain important counterfactual properties. Different theorists define these properties differently,

the properties go by many different names, and it is widely acknowledged that precisely formulating the properties is a delicate and tricky business (see, for example, Dretske 1971, Goldman 1979, Nozick 1981, DeRose 1995, Sosa 1999, Vogel 2000). But the general outlines of the counterfactual properties are clear enough to begin testing. The basic idea is that knowledge is, or essentially involves, “some variation of belief with the truth” (Nozick 1981: 209). The variation comes in two basic forms, which can be indicated by asking two simple questions. On the one hand, if things had been slightly different, *would* you still have believed the truth? On the other hand, if things had been slightly different, *might* you have believed falsely? For present purposes, I will call these properties *adherence* to truth and *resistance* to falsehood, respectively.<sup>66</sup>

Some claim that knowledge requires both adherence and resistance, whereas others claim that only one is required; I will ignore this dispute and instead investigate whether commonsense epistemology is implicitly committed to either principle and, if so, its potential connection to reliability. Some claim that knowledge requires even stronger or more elaborate forms of adherence or resistance; I will ignore these proposals and focus only on the basic forms. In order to test for a commitment to adherence, I will ask people whether an agent *would* still correctly identify the facts in slightly different circumstances. In order to test for a commitment to resistance, I will ask people whether an agent *might* misidentify the facts in slightly different circumstances. Experiment 8 tests adherence. Experiment 9 tests resistance.

---

<sup>66</sup> ‘Adherence’ is Nozick’s (1981: 212) term. I find it especially apt.

## Method

New participants ( $N = 142$ )<sup>67</sup> were randomly assigned to one of four groups in the same  $2$  (Source: Guess/Ability)  $\times$   $2$  (Percent: Thirty/Ninety) between-subjects design from Experiment 4, following the same procedures outlined for that experiment, except for one important difference. Participants recorded knowledge judgments in the same way (“knows” or “only thinks” multiplied by confidence, 1–6). This time, instead of an explanation probe, participants also answered an open adherence-counterfactual probe:

1. Suppose that the screen had flashed the word “bowl” instead. In that case, Carolyn \_\_\_\_\_ have reported “bowl”. (would/would not)

Participants then rated their confidence in their answer to the adherence probe on the same 1–6 scale.

## Results

The proto-realist hypothesis predicts, at the very least, that people will attribute knowledge only if they also attribute adherence. The proto-realist hypothesis also predicts that people will attribute adherence only for reliably produced belief. If the more ambitious claims of some realists are true, then when focusing on cases of true belief, we should also observe a strong positive correlation between knowledge judgments and adherence judgments, and perhaps we will even find evidence that adherence judgments cause or explain knowledge judgment.

---

<sup>67</sup> Aged 18-70 years, mean age = 31 years; 95% reporting English as a native language; 52 female. Twenty participants failed a comprehension question, but excluding these participants doesn't affect the pattern of results reported below.



To analyze the results, for each participant I calculated a knowledge score in the same way described in earlier experiments. I also calculated an *adherence score* similarly (“would” scored as +1, “would not” scored as -1, multiplied by reported confidence, 1–6). Higher adherence scores indicate greater adherence to the truth.

Assignment to condition significantly affected knowledge scores.<sup>68</sup> There was a trending main effect of Percent,<sup>69</sup> a large main effect of Source,<sup>70</sup> and no interaction.<sup>71</sup> Knowledge scores did not differ between Thirty Ability and Ninety Ability,<sup>72</sup> and they were significantly above the neutral midpoint (= 0) in each of those two conditions.<sup>73</sup> Knowledge scores were significantly higher in Ninety Guess than Thirty Guess,<sup>74</sup> but they were significantly below the midpoint in

---

68 ANOVA,  $F(3, 138) = 40.19$ ,  $p < .001$ ,  $\eta_p^2 = .466$ . An initial omnibus ANOVA revealed no main or interaction effects for participant age or gender: gender,  $F(1, 134) = 1.38$ ,  $p = .243$ ; age (two equal groups, cut-point = 28 years),  $F(1, 134) = 1.41$ ,  $p = .237$ .

69 ANOVA,  $F(1, 138) = 3.05$ ,  $p = .083$ ,  $\eta_p^2 = .022$ .

70 ANOVA,  $F(1, 138) = 117.30$ ,  $p < .001$ ,  $\eta_p^2 = .459$ .

71 ANOVA,  $F(1, 138) = 1.35$ ,  $p = .248$ .

72  $M = 3.23/3.58$ ,  $SD = 3.30/3.73$ . Independent samples t-test,  $t(69) = -0.424$ ,  $p = .673$ ,  $MD = -0.35 [-2.02, 1.31]$ .

73 One sample t-tests. Thirty Ability:  $t(34) = 5.79$ ,  $p < .001$ ,  $MD = 3.23 [2.10, 4.36]$ . Ninety Ability:  $t(35) = 5.77$ ,  $p < .001$ ,  $MD = 3.58 [2.32, 4.87]$ .

74  $M = -2.27/-4.03$ ,  $SD = 4.42/2.66$ ,  $t(59.8) = -2.05$ ,  $p = .045$ ,  $MD = -1.76 [-3.47, -0.04]$ . The magnitude of the difference in means was medium (Cohen's  $d = 0.52$ )

each of those two conditions.<sup>75</sup> Knowledge scores were significantly higher in Thirty Ability than Thirty Guess,<sup>76</sup> and they were higher in Ninety Ability than Ninety Guess.<sup>77</sup>

Assignment to condition significantly affected adherence scores.<sup>78</sup> There was a main effect of Source,<sup>79</sup> a trending main effect of Percent,<sup>80</sup> and no interaction.<sup>81</sup> Adherence scores were significantly higher in Ninety Ability than Thirty Ability,<sup>82</sup> and they were significantly above the neutral midpoint in each of those two conditions.<sup>83</sup> Adherence scores were no different in Ninety

---

75 One sample t-tests: Thirty Guess,  $t(33) = -8.84$ ,  $p < .001$ , MD = -4.03 [-4.96, -3.10]. Ninety Guess,  $t(36) = -3.12$ ,  $p = .004$ , MD = -2.27 [-3.74, -0.80].

76 Independent samples t-test,  $t(67) = -10.05$ ,  $p < .001$ , MD = 7.25 [5.82, 8.70]. The magnitude of the difference in means was extremely large (Cohen's  $d = 2.46$ ).

77 Independent samples t-test,  $t(69.6) = -6.12$ ,  $p < .001$ , MD = 5.85 [3.95, 7.76]. The magnitude of the difference in means was extremely large (Cohen's  $d = 1.46$ ).

78 ANOVA,  $F(3, 138) = 14.24$ ,  $p < .001$ ,  $\eta_p^2 = .236$ . An initial omnibus ANOVA revealed no main or interaction effects for participant age or gender.

79 ANOVA,  $F(1, 138) = 38.59$ ,  $p < .001$ ,  $\eta_p^2 = .236$ .

80 ANOVA,  $F(1, 138) = 2.93$ ,  $p = .089$ ,  $\eta_p^2 = .021$ .

81 ANOVA,  $F(1, 138) = 1.08$ ,  $p = .300$ .

82  $M = 2.89/4.75$ ,  $SD = 3.61/1.98$ . Independent samples t-test,  $t(52.4) = -2.69$ ,  $p = .010$ , MD = -1.86 [-3.26, -0.47]. The magnitude of the difference in means was large (Cohen's  $d = .74$ ).

83 One sample t-tests. Thirty Ability:  $t(34) = 4.73$ ,  $p < .001$ , MD = 2.89 [1.65, 4.13]. Ninety Ability:  $t(35) = 14.42$ ,  $p < .001$ , MD = 4.75 [4.08, 5.42].

Guess and Thirty Guess,<sup>84</sup> and they didn't differ from the midpoint in either condition.<sup>85</sup> Adherence scores were higher in Thirty Ability than Thirty Guess,<sup>86</sup> and they were higher in Ninety Ability than Ninety Guess.<sup>87</sup>

Overall, knowledge scores and adherence scores were moderately positively correlated.<sup>88</sup> However, unlike in Experiment 7 where explanation scores mediated the effect of condition on knowledge scores, in the present experiment adherence scores did not mediate the effect of con-

---

84 Thirty/Ninety Guess,  $M = -0.62/-0.16$ ,  $SD = 4.83/4.98$ . Independent samples t-test,  $t(69) = -0.39$ ,  $p = .698$ ,  $MD = 0.46 [-2.78, 1.87]$ .

85 One sample t-tests. Thirty Guess:  $t(33) = 0.745$ ,  $p = .462$ ,  $MD = -0.62 [-2.31, 1.07]$ . Ninety Guess:  $t(36) = -0.198$ ,  $p = .844$ ,  $MD = -0.16 [-1.82, 1.50]$ .

86 Independent samples t-test:  $t(61.05) = 3.4$ ,  $p = .001$ ,  $MD = 3.50 [1.44, 5.56]$ . The magnitude of the difference in means was large (Cohen's  $d = 0.88$ ).

87 Independent samples t-test:  $t(47.32) = 5.57$ ,  $MD = 4.91 [3.34, 6.69]$ . The magnitude of the difference in means was extremely large (Cohen's  $d = 1.62$ ).

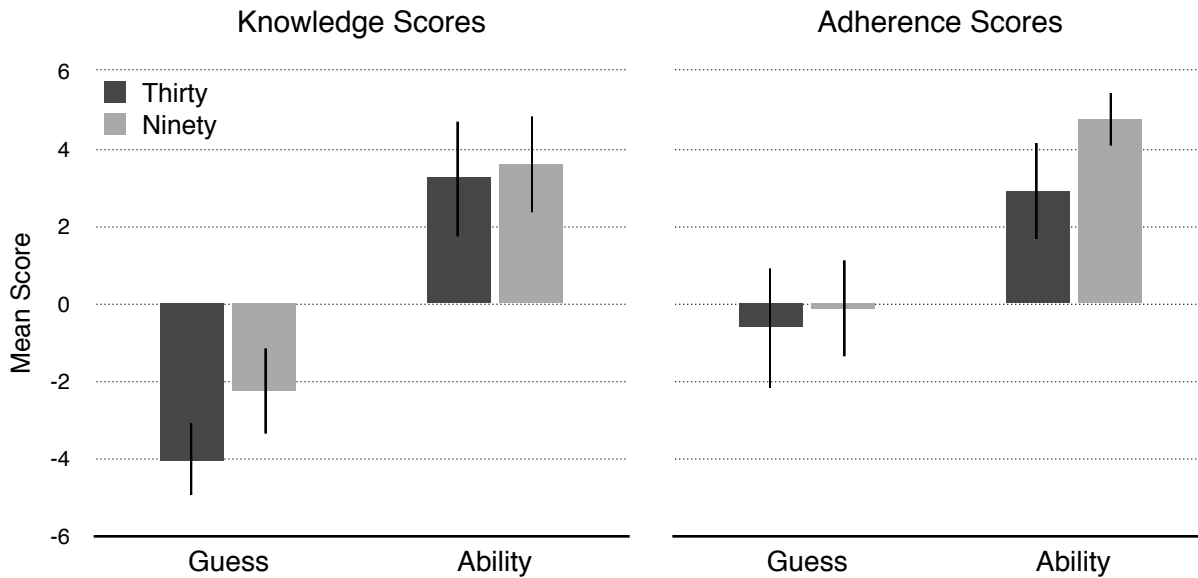
88 Pearson's correlation,  $r = .37$ ,  $N = 142$ ,  $p < .001$ .

dition on knowledge scores.<sup>89,90</sup>

---

89 As determined by two separate bootstrap mediation analyses. Comparing Ninety Ability and Ninety Guess, I entered Source as the independent variable, knowledge score as the outcome, and adherence score as the potential mediator: 95% CI for the indirect effect = -1.32 to 1.49; ratio of indirect to direct effect = .001, 95% CI for ratio = -0.28 to 0.29. Using the same procedure to compare Thirty Ability and Thirty Guess, there was a very weak indirect effect = 0.47, 95% CI = 0.02 to 1.38; ratio of indirect to direct effect = 0.06, 95% CI = 0.00 to 0.20. As a point of further information, in a multiple regression model with knowledge score as the outcome and Source, Percent, and adherence scores as predictors (model fit,  $R^2 = .464$ ), Source significantly predicted knowledge scores (Beta = .644,  $p < .001$ ), but Percent (Beta = .101,  $p = .111$ ) and adherence scores (Beta = .060,  $p = .402$ ) did not.

90 A causal search with the Greedy Equivalence Search (GES) algorithm returned a model in which Source caused both knowledge judgments and adherence judgments, Percent caused neither judgment, and there was no causal connection between knowledge judgments and adherence judgments. The model fit the data well,  $\chi^2(4) = 6.74$ ,  $p = .15$ , BIC = -13.08. The causal search was conducted with Tetrad 5 (<http://www.phil.cmu.edu/projects/tetrad/>). GES considers all possible models available given the different variables. Each variable is treated as a node. GES assigns an information score to the model in which all the nodes are disconnected — the “null model.” GES then evaluates adding causal arrows — “edges” — between the nodes. GES adds edges that best improve the model’s information score, if such edges exist, until adding more edges does not improve the information score. At this point, GES evaluates whether deleting any edges will further improve the information score, and it deletes any such edges. See Rose & Nichols 2013 and Rose, Livengood, Sytsma & Machery 2012 for prior applications of GES to philosophical debates, as well as the literature cited therein for further details on the algorithm.



**Fig. 8.** Experiment 8: Left panel: mean knowledge scores, derived for each participant by multiplying their dichotomous answer (scored +1 for ‘knows’, -1 for ‘only thinks’) by their reported confidence (1 = “not at all confident” to 6 = “completely confident”). Right panel: mean adherence scores, derived for each participant by multiplying their dichotomous answer (scored +1 for ‘would’, -1 for ‘would not’) by their reported confidence. Scales ran from -6 to +6. Error bars represent 95% confidence intervals.

## Discussion

The results are consistent with the prediction that people share an “adherence intuition,” whereby they tend to attribute knowledge only if they also tend to attribute adherence. In other words, the results are consistent with the hypothesis that knowledge, as it’s ordinarily understood, implies adherence. But the results also show that reliability does not explain the relationship between knowledge and adherence. That is, assuming that adherence is a principle of folk epistemology, reliability is not essential to it: people think that *unreliably* produced belief exhibits (enough) adherence to count as knowledge. Moreover, although people were more likely to agree that a reliably produced belief is adherent, this increase in adherence judgments did not increase rates of

knowledge attribution. Additionally, although knowledge and adherence scores were moderately positively correlated, I found no evidence that people make knowledge judgments because they make adherence judgments.

The results once again replicate the main findings that people readily attribute unreliable knowledge, and that vast differences in reliability don't affect knowledge judgments.

One concern with this experiment pertains to how effectively participants integrate relevant information about reliability when making counterfactual judgments. Explicit counterfactual judgments can be cognitively demanding, which could contribute to this. The next and final experiment investigates the other potential counterfactual property of knowledge, resistance, and includes a check to determine whether people effectively integrate information about reliability when making counterfactual judgments.

## **Experiment 9**

### **Method**

New participants (N = 146)<sup>91</sup> were randomly assigned to one of four groups in the same 2 (Source: Guess/Ability) × 2 (Percent: Thirty/Ninety) between-subjects design from Experiment 8, following the same procedures outlined for that experiment, except for one important difference. Participants recorded knowledge judgments in the same way (“knows” or “only thinks”

---

91 Aged 18-63 years, mean age = 31 years; 95% reporting English as a native language; 57 female.

multiplied by confidence, 1–6). This time, instead of an adherence probe, participants also responded to a *resistance*-counterfactual probe:

1. Suppose that instead of flashing the word “corn”, the screen had flashed some other, different word instead. In such a case, might Carolyn make a mistake about what the word was? (No/Yes)

Participants then rated their confidence in their answer to the resistance probe on the same 1–6 scale. Finally, participants then went to a new screen to answer a counterfactual manipulation check.

2. Again, suppose that instead of flashing the word “corn”, the screen had flashed some other, different word instead. How likely is it that Carolyn would correctly identify that other word?

Responses to this question were collected on a standard 6-point Likert scale, anchored with “very unlikely” (= 1), “unlikely,” “somewhat unlikely,” “somewhat likely,” “likely,” and “very

likely” (= 6).<sup>92</sup>

## Results

Preliminary analysis revealed that people effectively integrated information about reliability when making counterfactual judgments. Whether Carolyn was guessing or exercising her ability, when she was highly reliable, participants answered that it was significantly more likely that she

---

92 In an attempt to detect subtler aspects of the potential relationship between reliability and knowledge judgments, I also collected reaction time data in this experiment. More specifically, I collected information on how long it took people to register their knowledge judgments. (The timing mechanism was invisible to participants and they were not told that they were being timed.) It could turn out that although people attribute reliable and unreliable knowledge at equal rates, reliability facilitates more rapid knowledge attribution, or unreliability facilitates more rapid knowledge denial. If this were true, then it might suggest that reliability is associated with prototypical knowledge, whereas unreliability is not. In this way, reliability might be importantly conceptually connected to knowledge without being a necessary condition on knowledge. This could explain why philosophers are so easily tempted by the false view that, as a matter of conceptual necessity, knowledge requires reliability. The explanation would be that philosophers have correctly sensed a conceptual connection but misidentified its nature. However, I found no evidence for this hypothesis, because it turned out that people recorded their knowledge judgments faster when the source of belief was more unreliable.



would judge correctly in slightly different counterfactual circumstances.<sup>93</sup>

The proto-reliabilist hypothesis predicts, at the very least, that people will attribute knowledge only if they also attribute resistance. The proto-reliabilist hypothesis also predicts that people will attribute resistance only for reliably produced belief. If the more ambitious claims of some reliabilists are true, then when focusing on cases of true belief, we should also observe a strong positive correlation between knowledge judgments and resistance judgments, and perhaps we will even find evidence that resistance judgments cause or explain knowledge judgments.

To analyze the results, for each participant I calculated a knowledge score in the same way described in earlier experiments. I also calculated a *resistance score* similarly (“No” scored as +1, “Yes” scored as -1, multiplied by reported confidence, 1–6). Higher resistance scores indicate greater resistance to making a mistake.

The proto-reliabilist predictions turned out false. Assignment to condition significantly af-

---

93 Independent samples t-test. Comparing Ninety Guess and Thirty Guess:  $M = 4.37/3.59$ ,  $SD = 1.10/1.07$ ,  $t(75) = 3.15$ ,  $p = .002$ ,  $MD = 0.78 [0.29, 1.27]$ . Comparing Ninety Ability and Thirty Ability:  $4.97/3.76$ ,  $SD = 1.04/0.99$ ,  $t(67) = 4.94$ ,  $p < .001$ ,  $MD = 1.21 [0.72, 1.70]$ . One sample t-tests revealed that likelihood scores were significantly above the theoretical midpoint (= 3.5) in Ninety Guess,  $t(37) = 4.86$ ,  $p < .001$ , and in Ninety Ability,  $t(34) = 8.35$ ,  $p < .001$ . By contrast, likelihood scores did not differ from the midpoint in either Thirty Guess,  $t(38) = 0.52$ ,  $p = .603$ , or Thirty Ability,  $t(33) = 1.57$ ,  $p = .127$ .

fect knowledge scores.<sup>94</sup> There was a large main effect of Source,<sup>95</sup> no effect of Percent,<sup>96</sup> and no interaction.<sup>97</sup> Knowledge scores did not differ between Thirty Ability and Ninety Ability,<sup>98</sup> and they were significantly above the neutral midpoint (= 0) in each of those two conditions.<sup>99</sup> Knowledge scores did not differ between Thirty Guess and Ninety Guess,<sup>100</sup> and they were significantly below the midpoint in each of those two conditions.<sup>101</sup> Knowledge scores were significantly higher in Thirty Ability than Thirty Guess,<sup>102</sup> and they were higher in Ninety Ability than

---

94 ANOVA,  $F(3, 142) = 28.73$ ,  $p < .001$ ,  $\eta_p^2 = .378$ . An initial omnibus ANOVA revealed no main or interaction effects for participant age or gender: gender,  $F(1, 130) = 2.05$ ,  $p = .155$ ; age (two equal groups, cut-point = 29 years),  $F(1, 130) = 2.18$ ,  $p = .142$ .

95 ANOVA,  $F(1, 142) = 83.42$ ,  $p < .001$ ,  $\eta_p^2 = .370$ .

96 ANOVA,  $F(1, 142) = 2.01$ ,  $p = .159$ .

97 ANOVA,  $F(1, 142) = 0.313$ ,  $p = .577$ .

98 Thirty/Ninety Ability,  $M = 2.38/3.94$ ,  $SD = 4.14/3.81$ . Independent samples t-test,  $t(67) = -0.585$ ,  $p = .560$ ,  $MD = -0.56 [-2.47, 1.35]$ .

99 One sample t-tests. Thirty Ability,  $t(33) = 3.354$ ,  $p = .002$ ,  $MD = 2.38 [0.93, 3.83]$ . Ninety Ability,  $t(34) = 4.57$ ,  $p < .001$ ,  $MD = 2.94 [1.63, 4.25]$ .

100 Thirty/Ninety Guess,  $M = -3.95/-2.66$ ,  $SD = 3.39/4.37$ . Independent samples t-test,  $t(69.8) = -1.45$ ,  $p = .153$ ,  $MD = -1.29 [-3.07, 0.49]$ .

101 One sample t-tests. Thirty Guess,  $t(38) = -7.26$ ,  $p < .001$ ,  $MD = -3.95 [-5.05, -2.85]$ . Ninety Guess,  $t(37) = -3.75$ ,  $p = .001$ ,  $MD = -2.66 [-4.09, -1.22]$ .

102 Independent samples t-test,  $t(63.93) = 7.08$ ,  $p < .001$ ,  $MD = 6.33 [4.54, 8.12]$ . The magnitude of the difference in means was extremely large (Cohen's  $d = 1.77$ ).

Ninety Guess.<sup>103</sup>

Assignment to condition significantly affected resistance scores.<sup>104</sup> There were small main effects of Source<sup>105</sup> and Percent<sup>106</sup> and no interaction.<sup>107</sup> Resistance scores were higher in Ninety Ability than in Thirty Ability,<sup>108</sup> and they were no different between Ninety Guess and Thirty Guess.<sup>109</sup> Resistance scores were below the midpoint in all four conditions, significantly so for Thirty Ability, Thirty Guess, and Ninety Guess.<sup>110</sup>

Overall, knowledge scores and resistance scores were moderately positively correlated.<sup>111</sup>

---

103 Independent samples t-test,  $t(71) = 5.81$ ,  $p < .001$ , MD = 5.60 [3.68, 7.52]. The magnitude of the difference in means was extremely large (Cohen's  $d = 1.38$ ).

104 ANOVA,  $F(3, 142) = 4.318$ ,  $p = .006$ ,  $\eta_p^2 = .084$ .

105 ANOVA,  $F(1, 142) = 6.82$ ,  $p = .010$ ,  $\eta_p^2 = .046$ .

106 ANOVA,  $F(1, 142) = 4.51$ ,  $p = .036$ ,  $\eta_p^2 = .031$ .

107 ANOVA,  $F(1, 142) = 1.77$ ,  $p = .186$ .

108 Ninety/Thirty Ability,  $M = -0.54/-2.76$ ,  $SD = 4.84/4.16$ . Independent samples t-test,  $t(66.02) = 2.05$ ,  $p = .045$ , MD = 2.22 [0.06, 4.39]. The magnitude of the difference in means was medium (Cohen's  $d = 0.50$ ).

109 Ninety/Thirty Guess,  $M = -3.08/-3.59$ ,  $SD = 3.27/3.16$ . Independent samples t-test,  $t(75) = -0.70$ ,  $p = .488$ , MD = 0.51 [-0.95, 1.97].

110 One sample t-tests. Thirty Guess,  $t(38) = -7.09$ ,  $p < .001$ , MD = -3.59 [-4.61, -2.57]. Thirty Ability,  $t(33) = -3.88$ ,  $p < .001$ , MD = -2.76 [-4.22, -1.31]. Ninety Guess,  $t(37) = -5.81$ ,  $p < .001$ , MD = -3.08 [-4.15, -2.01]. Ninety Ability,  $t(34) = -0.66$ ,  $p = .512$ , MD = -0.54 [-2.21, 1.12].

111 Pearson's correlation,  $r = .422$ ,  $N = 146$ ,  $p < .001$ .

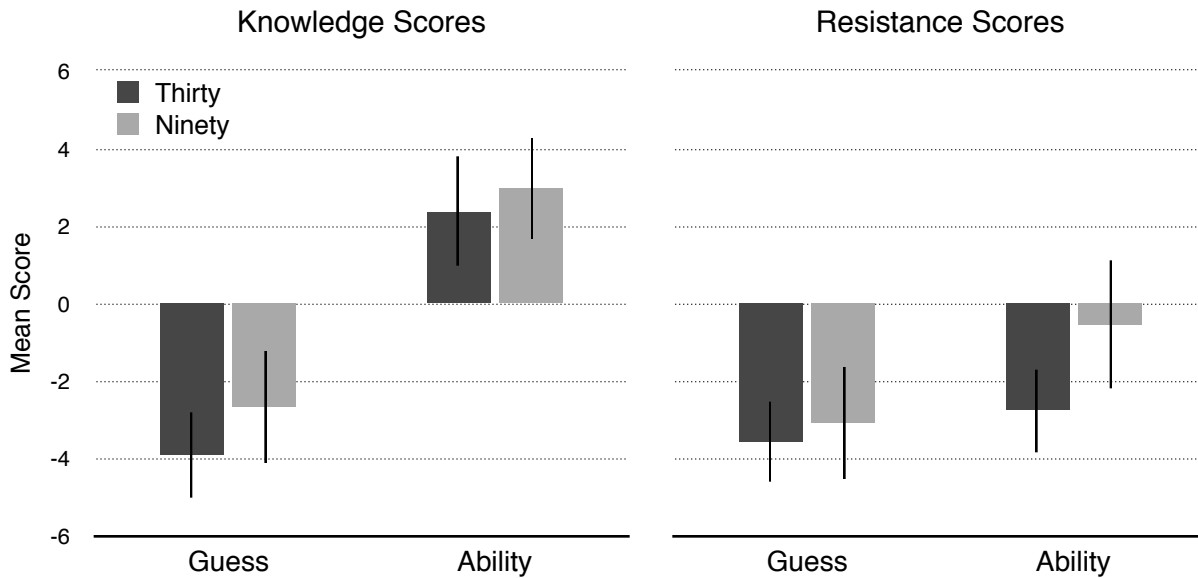
Critically, knowledge scores were significantly higher than resistance scores in Thirty Ability<sup>112</sup> and Ninety Ability.<sup>113</sup> Also critically, even though knowledge scores were much higher in Thirty Ability than Thirty Guess, resistance scores in the two conditions didn't differ.<sup>114</sup>

---

112 Paired sample t-test,  $t(33) = 6.54$ ,  $p < .001$ ,  $MD = 5.14$  [3.55, 6.75]. The magnitude of the difference in means was very large (Cohen's  $d = 1.24$ ).

113 Paired sample t-test,  $t(34) = 4.55$ ,  $p < .001$ ,  $MD = 3.49$  [1.93, 5.04]. The magnitude of the difference in means was large (Cohen's  $d = 0.80$ ).

114 Independent samples t-test,  $t(71) = 0.96$ ,  $p = .340$ ,  $MD = 0.83$  [-0.89, 2.54]. A causal search with the Greedy Equivalence Search (GES) algorithm returned a model in which Source was the sole cause of knowledge judgments, and knowledge judgments were the sole cause of resistance judgments. Resistance judgments and Percent conspired to cause counterfactual likelihood judgments (i.e. response to the manipulation check). The model fit the data well,  $\chi^2(6) = 9.55$ ,  $p = .15$ ,  $BIC = -20.35$ .



**Fig. 9.** Experiment 9: Left panel: mean knowledge scores, derived for each participant by multiplying their dichotomous answer (scored +1 for ‘knows’, -1 for ‘only thinks’) by their reported confidence (1 = “not at all confident” to 6 = “completely confident”). Right panel: mean resistance scores, derived for each participant by multiplying their dichotomous answer (scored +1 for ‘No’, -1 for ‘Yes’) by their reported confidence. Scales ran from -6 to +6. Error bars represent bootstrapped 95% confidence intervals.

## Discussion

The results show that there probably is no “resistance intuition” for reliabilism to explain. The results are inconsistent with the prediction that people will attribute knowledge only if they attribute resistance to error. In other words, the results undermine the hypothesis that knowledge ordinarily understood implies resistance. People were very willing to attribute knowledge, even though they were strongly inclined to agree that the agent might have been wrong in very similar counterfactual circumstances. I also observed vast differences in rates of knowledge attribution between cases that did not differ significantly in their extremely low rates of resistance attribution. This all happened most pointedly in cases of unreliably produced knowledge.

Whatever level of resistance is required for knowledge, reliability is not essential to it. People are happy to agree that unreliably produced belief exhibits (enough) resistance to count as knowledge. And although people were more likely to agree that a reliably produced belief is resistant, this increase in resistance judgments did not increase knowledge attribution. Moreover, not even highly reliably produced belief was viewed as resistant, even though the vast majority of people judged it to be knowledge. Finally, the results once again replicated the main findings that people readily attribute unreliable knowledge, and that vast differences in reliability don't affect knowledge judgments.

## **Conclusion**

The view that knowledge must be reliably produced, knowledge reliabilism, is extremely popular in contemporary philosophy. But its popularity is due to inconclusive arguments and recently it has been challenged on theoretical grounds. Some have speculated that reliabilism draws support from the ordinary concept of knowledge and patterns in ordinary judgments about knowledge. That is, some have speculated that “commonsense” or “folk” epistemology is implicitly a form of proto-reliabilism and that reliabilist epistemology merely “explicates the ordinary standards” that guide knowledge judgments. This paper reported nine experiments designed to test the proto-reliabilist hypothesis.

The results show that the hypothesis is false. Folk epistemology fully embraces the possibility of unreliable knowledge. Multiple times I observed knowledge attributions reach 80-90% in

cases of unreliably produced true belief. Participants consistently attributed knowledge to reliable and unreliable believers at similar rates. Participants attributed knowledge despite actively classifying the knower as unreliable, and they overwhelmingly declined to attribute knowledge in closely matched controls. More generally, although people understand and process explicit information about reliability, they do not seem to consult this information when making knowledge judgments. I observed the same basic pattern using different dependent measures (scaled and dichotomous), across different cover stories (computerized tests and ordinary tasks), for male and female believers, and for different faculties (vision and memory). None of this would happen if reliabilism was in sync with commonsense epistemology.

Reliabilists have also mischaracterized the “fundamental” intuitive relationship between knowledge and luck (e.g. Heller 1999, Lewis 1996, Dretske 1981, Pritchard 2005; see also Hetherington 2013). Almost no one thinks that knowledge is attributable to luck rather than ability, and it appears that people’s judgments about luck and ability guide their knowledge judgments, so commonsense does endorse an “anti-luck” intuition (Pritchard 2013). But most people do not share the intuition that *reliability* specifically is what rules out epistemically problematic luck.<sup>115</sup> Instead, a better candidate for capturing the intuitive relationship is that ability, reliable or unreli-

---

115 Here I use “reliable” as knowledge reliabilists do, namely, indicating success on the (vast) majority of trials (see the Introduction for references). For the record, while I don’t begrudge researchers adapting terminology to their own purposes, I question whether this corresponds to the ordinary meaning of “reliable” or is even the sense of “reliable” most relevant to assessing knowledge and other accomplishments. For more discussion, see Turri 2015a.

able, is what rules out luck.<sup>116</sup> At the same time, a nontrivial minority denied knowledge despite attributing true belief to ability, albeit a fairly unreliable one that produces false beliefs seventy percent of the time. This suggests that there are important individual differences in how much unreliability is viewed as consistent with knowledge or, perhaps more fundamentally, having an ability. Future work could investigate what these differences are and whether philosophers tend to have traits that predict greater intolerance of unreliability (compare Feltz & Cokely 2012). The present results suggest that neither gender nor age contributes to this intolerance in a consistent or meaningful way.

Reliabilists have also mischaracterized the intuitive relationship among reliability, knowledge, and counterfactual judgments. I studied two types of counterfactual judgments: judgments of *adherence to the truth* and *resistance to error* in close counterfactual circumstances. On the one hand, while the present results are consistent with the hypothesis that knowledge implies adherence in folk epistemology, I found that reliability is not essential to adherence. Instead, forming a true belief through unreliable ability brings with it enough adherence for knowledge. Moreover, it does not appear that people make knowledge judgments because of their adherence judgments — that is, people do not base their knowledge judgments on adherence judgments. However, I emphasize that this causal claim is preliminary and more work is needed before drawing firm conclusions.

---

116 As argued elsewhere on theoretical grounds (Turri 2012, Turri 2015a). Some philosophers seem to have recognized that ability is what rules out luck, but they pair this with the mistaken view that ability requires reliability (e.g. Sosa 2007, Greco 2010, Pritchard 2012).



On the other hand, I found evidence that resistance is not viewed as a requirement of knowledge. Instead, people attributed knowledge despite recognizing that the knower easily might have made a mistake. Prior research has found that people are willing to attribute knowledge while acknowledging that “there is at least some chance, no matter how small” that the person was wrong (Turri & Friedman 2014). The present findings on “might”-counterfactuals, and unreliably produced knowledge more generally, deepen our appreciation for the depth of fallibilist sentiment in folk epistemology. At this point, any form of infallibilism about knowledge should be viewed as radically revisionary and, consequently, bestow significant advantages before being taken seriously. This pertains to views that require some form of unrestricted infallibility (as suggested, for example, by Bonjour 2010 and Unger 1975), as well as “contextualist” or “relevant alternative” approaches that require infallibility relative to some restricted set of possibilities (as suggested by Dretske 1970, Stine 1976, Goldman 1976, Unger 1984, Cohen 1988, DeRose 1992, DeRose 1995, Lewis 1996, Cohen 2013).

I found two small pieces of evidence of reliabilist sentiment in folk epistemology, but this evidence raises more questions than it answers. On the one hand, in Experiment 2, reliability significantly boosted knowledge attributions in cases of *false* belief to chance rates (~40% compared to ~10% in cases of unreliably formed false belief). This puzzling finding could have been an anomaly; it did not replicate in Experiment 3-5. On the other hand, in Experiment 8, reliability significantly boosted knowledge attributions in cases where the person *just guesses* the correct answer (~30% compared to ~10% in cases where the person who just guesses the correct answer is unreliable). Again, this puzzling finding could have been an anomaly; it did not occur in Ex-

periments 7 or 9. But even if these small and counterintuitive results turn out to be robust, in no way do they vindicate the proto-reliabilist hypothesis. Indeed, reliabilists might risk some embarrassment by touting them.

Some philosophers have wondered about our tendency in ordinary settings to count someone as knowing when all we know is that he has a true belief. For instance, some authors note that knowledge is “epistemically richer” than true belief and proceed to deny that our tendencies reveal that we ordinarily think of knowledge as mere true belief (Sripada & Stanley 2012, p. 7, discussing an example due to Hawthorne 2000). These authors hypothesize that “the folk are so easily led to mistake . . . true belief for knowledge” because, in many contexts, “all we care about” is that the “known” claim is true. In other words, lack of concern leads the folk to grant knowledge. But the findings from Experiment 7 suggest a completely different and more charitable explanation of any such tendency: when someone forms a true belief and has a cognitive ability relevant to acquiring the truth, the default assumption is that she formed the true belief through ability and thus knows.

The results also suggest an explanation for a recent development that has surprised many philosophers. For decades philosophers claimed that, intuitively, agents in “fake barn” cases lack knowledge (e.g. Goldman 1976; Sosa 1991: 238-9; Pritchard 2005: 161-2; Kvanvig 2008: 274). In a fake barn case, the agent (“Henry” in the original) perceives an object (a barn in the original) and this perceptual relation remains intact throughout. It is then revealed that the agent currently inhabits an environment filled with numerous fakes (“papier-mâché facsimiles of barns” in the original), which look like real ones. In an effort to explain why we are “strongly inclined” to

deny knowledge in such cases, philosophers appeal to the agent's high likelihood of forming a false belief in the context (Goldman 1976: 773; Pritchard 2014: ch. 6; Plantinga 1993: 33).

But such "explanations" are idle if the tendency does not exist. And recent research has shown that in fact the tendency does not exist. Instead, people tend to attribute knowledge in fake-barn cases (Colaco, Buckwalter, Stich & Machery 2014). More generally in cases where an agent perceives something that makes her belief true and this perceptual relation remains intact, people overwhelmingly attribute knowledge, even when lookalikes pose a salient threat to the truth of the agent's belief (Turri, Buckwalter & Blouw 2014; see also Turri in press b; Turri in press c). Indeed, commonsense views these similarly to *paradigm instances* of knowledge, with rates of knowledge attribution often exceeding 80% — which closely resembles the rate of unreliable knowledge attribution repeatedly observed in the present experiments. The present results thus suggest that philosophers were wrong about fake barn cases for the same reason that they were wrong about reliabilism: they assumed that knowledge requires reliability.

There is a view that coherently and elegantly unites all of these findings: *abilism*. Abilism defines knowledge as true belief manifesting the agent's cognitive ability or powers.<sup>117</sup> Crucially, cognitive abilities can be unreliable in the sense that they needn't produce true beliefs more often than not. Perhaps they need only produce true beliefs at rates exceeding chance. In fake barn cases and in the cases studied here, the agent is best interpreted as forming or retaining a true belief

---

117 This is an approximate and simplified statement of the view. Fully spelled out, knowledge is *approximately* true *thin* belief manifesting cognitive power. For more on the motivation for these refinements, see (Turri 2015b).

through cognitive ability (vision or memory). Thus he is best interpreted as knowing, which explains the extremely high rates of knowledge attribution. Moreover, as discussed in the Introduction, abilism can explain Goldman's initial observations intended to motivate reliabilism. (For further ways that abilism co-opts reliabilism's explanatory exploits, see Turri 2012). Additionally, abilism is well suited to explain the present findings on epistemic luck and the enormous difference in knowledge judgments between the Ability and Guess conditions from Experiment 7, and the large differences between the Thirty Ability and Thirty Guess conditions in Experiments 8 and 9. Abilism is also fully consistent with the findings on the relationship between knowledge and counterfactual judgments.

Before concluding, I should note a potentially important limitation of the present findings and the recent findings on fake barn cases. All the cases tested involve beliefs that are naturally interpreted as *non-inferential*, relatively "immediate" beliefs based on perception and memory (Goldman 2008). But cognitive scientists have recently discovered many important qualitative differences in how inferential and non-inferential beliefs are evaluated (Turri & Friedman 2014; Turri 2015c; Turri 2015d; Zamir, Rito & Teichman in press; see also Pillow 1989, Sodian & Wimmer 1987). Further work could investigate whether *inferential* knowledge specifically requires reliability or interacts differently with judgments about luck. Relatedly, further work could also investigate whether the bar is set much higher for the possession of inferential cognitive abilities (discovering truths through reasoning) than for non-inferential cognitive abilities (e.g. perceptually detecting truths and retaining them through memory). Perhaps the intuitions motivating reliabilism are keyed to inferential knowledge or cognitive abilities, and philosophers mis-

takenly generalized this to cover all knowledge or cognitive abilities.

Regardless of what might be discovered about inferential knowledge specifically, the present findings constitute another exhibit in the mounting case against knowledge reliabilism, which is currently endorsed by almost all epistemologists. We have passed the point where it is reasonable to assume that reliabilism is true. Indeed, quite the opposite — moving forward, we should assume that reliabilism is false. Potent theoretical and empirical considerations have converged to support the conclusion that unreliable knowledge is not only possible but actual and widely recognized as perfectly ordinary. A new day has dawned. It's time for the discipline to wake from its reliabilist slumber.

*Acknowledgments* — For helpful feedback and discussion, I thank Wesley Buckwalter, A.Y. Darling, Sandra DeVries, Ori Friedman, John Greco, Ashley Keefner, Josh Knobe, Janet Michaud, Blake Myers, Shaylene Nancekivell, Luis Rosa, David Rose, Jamie Sewell, Angelo Turri, and Sara Weaver. Thanks also to an audience at the 2014 Buffalo Experimental Philosophy Conference and participants in my 2014-15 seminar on experimental philosophy at the University of Waterloo. This research was supported by the Social Sciences and Humanities Research Council of Canada and an Early Researcher Award from the Ontario Ministry of Economic Development and Innovation.

*Word count:* 9597 (w/footnotes and references: 15744)

## References

- Alexander, J. (2012). *Experimental Philosophy*. Cambridge: Polity.
- Alston, W. P. (1988). An internalist externalism. *Synthese*, 74(3), 265–283.
- Alston, W. P. (1993). *The reliability of sense perception*. Ithaca, N.Y. : Cornell University Press.
- Alston, W. P. (1995). Ho to think about reliability. *Philosophical Topics*, 23(1), 1–29.
- Armstrong, D. M. (1973). *Belief, truth and knowledge*. Cambridge: Cambridge University Press.
- Austin, J. L. (1956). A plea for excuses. *Proceedings of the Aristotelian Society*, 57, 1–30.
- Becker, K. (2006). Reliabilism and safety. *Metaphilosophy*, 37(5), 691–704.
- Beebe, J. R. (2012). Experimental epistemology. In A. Cullison (Ed.), *Continuum companion to epistemology* (pp. 248–269). Continuum.
- Beebe, J. R. (Ed.). (2014). *Advances in experimental epistemology*. Bloomsbury Academic.
- BonJour, L. (2002). *Epistemology: classic problems and contemporary responses*. Rowman & Littlefield.
- BonJour, L. (2010). The myth of knowledge. *Philosophical Perspectives*, 24(1), 57–83.
- Cohen, S. (1984). Justification and truth. *Philosophical Studies*, 46(3), 279–295.
- Cohen, S. (2002). Basic knowledge and the problem of easy knowledge. *Philosophy and Phenomenological Research*.
- Cohen, S. (2013). Contextualism defended. In M. Steup, J. Turri, & E. Sosa (Eds.), *Contemporary debates in epistemology* (2nd ed., pp. 69–75). Malden, Mass.: Wiley-Blackwell.
- Colaco, D., Buckwalter, W., Stich, S., & Machery, E. (2014). Epistemic intuitions in fake-barn

- thought experiments. *Episteme*, 11(02), 199–212. doi:10.1017/epi.2014.7
- Comesaña, J. (2010). Evidentialist reliabilism. *Nous*, 44(4), 571–600.
- Comesaña, J. (2002). The diagonal and the demon. *Philosophical Studies*, 110(2), 249–266.
- Craig, E. (1990). *Knowledge and the state of nature: an essay in conceptual synthesis*. Oxford: Oxford University Press.
- Cullen, S. (2010). Survey-Driven Romanticism. *Review of Philosophy and Psychology*, 1(2), 275–296. doi:10.1007/s13164-009-0016-1
- DeRose, K. (1992). Contextualism and knowledge attributions. *Philosophy and Phenomenological Research*, 52(4), 913–929.
- DeRose, K. (1995). Solving the skeptical problem. *The Philosophical Review*, 104(1), 1–52.
- Dretske, F. (1971). Conclusive reasons. *Australasian Journal of Philosophy*.
- Dretske, F. I. (1970). Epistemic operators. *The Journal of Philosophy*, 67(24), 1007–1023.
- Dretske, F. I. (1981). *Knowledge and the flow of information*. Cambridge, Mass.: MIT Press.
- Fantl, J., & McGrath, M. (2009). *Knowledge in an uncertain world*. Oxford University Press.
- Feldman, R. (2003). *Epistemology*. Upper Saddle River, NJ: Prentice Hall.
- Feldman, R. (2013). Justification is internal. In M. Steup, J. Turri, & E. Sosa (Eds.), *Contemporary debates in epistemology* (2nd ed.). Malden, Mass.: Wiley-Blackwell.
- Feltz, A., & Cokely, E. T. (2012). The Philosophical Personality Argument. *Philosophical Studies*, 161(2), 227–246. doi:10.1007/s11098-011-9731-4
- Foley, R. (1985). What's wrong with reliabilism? *The Monist*, 68(2), 188–202.
- Fumerton, R. (2006). *Epistemology*. Malden, MA: Blackwell.

- Goldman, A. I. (1976). Discrimination and perceptual knowledge. *Journal of Philosophy*, 73(20), 771–791.
- Goldman, A. I. (1979). What is justified belief? In G. Pappas (Ed.), *Justification and knowledge*. Dordrecht: Reidel.
- Goldman, A. I. (1986). *Epistemology and Cognition*. Harvard University Press.
- Goldman, A. I. (1993). Epistemic folkways and scientific epistemology. *Philosophical Issues*, 3, 271–285.
- Goldman, A. I. (2008). Immediate justification and process reliabilism. In Q. Smith (Ed.), *Epistemology: new essays* (pp. 63–82). Oxford: Oxford University Press.
- Goldman, A. I. (2012). *Reliabilism and contemporary epistemology*. Oxford: Oxford University Press.
- Goldman, A. I., & Olsson, E. J. (2009). Reliabilism and the Value of Knowledge. In A. Millar, A. Haddock, & D. Pritchard (Eds.), *Epistemic value* (pp. 19–41). Oxford: Oxford University Press.
- Greco, J. (2000). *Putting skeptics in their place : the nature of skeptical arguments and their role in philosophical inquiry*. Cambridge: Cambridge University Press.
- Greco, J. (2010). *Achieving knowledge: A virtue-theoretic account of epistemic normativity*. Cambridge: Cambridge University Press.
- Greco, J. (2013). Justification is not internal. In M. Steup, J. Turri, & E. Sosa (Eds.), *Contemporary debates in epistemology* (2nd ed.). Malden, Mass.: Wiley-Blackwell.
- Hawthorne, J. (2000). Implicit Belief and A Priori Knowledge. *The Southern Journal of Philoso-*



- phy*, 38(S1), 191–210. doi:10.1111/j.2041-6962.2000.tb00937.x
- Hawthorne, J. (2004). *Knowledge and lotteries*. Oxford: Oxford University Press.
- Hayes, A. F. (2013). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. New York: Guilford Press.
- Heller, M. (1995). The simple solution to the problem of generality. *Nous*, 29(4), 501–515.
- Heller, M. (1999). The proper role for contextualism in an anti-luck epistemology. *Philosophical Perspectives*, 13, 115–129. doi:10.1111/0029-4624.33.s13.5
- Hetherington, S. (2013). There can be lucky knowledge. In M. Steup, J. Turri, & E. Sosa (Eds.), *Contemporary debates in epistemology*. (2nd ed.). Malden, Mass.: Wiley-Blackwell.
- Hill, C. S. (1996). Process reliabilism and cartesian scepticism. *Philosophy and Phenomenological Research*, 56(3), 567–581.
- Kornblith, H. (2002). *Knowledge and its place in nature*. Oxford: Oxford University Press.
- Kvanvig, J. (2008). Epistemic luck. *Philosophy and Phenomenological Research*, 77(1), 272–281. doi:10.1111/j.1933-1592.2008.00187.x
- Lehrer, K. (1990). *Theory of knowledge*. Boulder: Westview Press.
- Lehrer, K. (1997). *Self-trust*. Oxford University Press.
- Lehrer, K., & Cohen, S. (1983). Justification, truth, and coherence. *Synthese*, 55(2), 191–207.
- Lewis, D. (1996). Elusive knowledge. *Australasian Journal of Philosophy*, 74(4), 549–567.
- Nozick, R. (1981). *Philosophical explanations*. Cambridge, Mass: Harvard University Press.
- Pinillos, N. Á. (2011). Some Recent Work in Experimental Epistemology. *Philosophy Compass*, 6(10), 675–688. doi:10.1111/j.1747-9991.2011.00440.x

- Plantinga, A. (1993). *Warrant and proper function*. Oxford: Oxford University Press.
- Pritchard, D. (2005). *Epistemic luck*. New York : Oxford University Press.
- Pritchard, D. (2009). Apt performance and epistemic value. *Philosophical Studies*, 143(3), 407–416.
- Pritchard, D. (2012). Anti-luck virtue epistemology. *Journal of Philosophy*, 109(3), 247–279.
- Pritchard, D. (2013). Knowledge cannot be lucky. In M. Steup, J. Turri, & E. Sosa (Eds.), *Contemporary debates in epistemology* (2nd ed., pp. 152–164). Malden, Mass.: Wiley-Blackwell.
- Pritchard, D. (2014). *What is this thing called knowledge?* (3rd ed.). New York: Routledge.
- Riggs, W. D. (2002). Reliability and the value of knowledge. *Philosophy and Phenomenological Research*, 64(1), 79–96.
- Rose, D., & Nichols, S. (2013). The Lesson of Bypassing. *Review of Philosophy and Psychology*, 4, 599–619.
- Rose, D., Livengood, J., Sytsma, J., & Machery, E. (2012). Deep trouble for the deep self. *Philosophical Psychology*, 25(5), 629–646.
- Roush, S. (2005). *Tracking truth*. Oxford: Oxford University Press.
- Russell, B. (2012). *Human knowledge: its scope and limits*. New York: Routledge.
- Sartwell, C. (1992). Why knowledge is merely true belief. *The Journal of Philosophy*, 89(4), 167–180.
- Sellars, W. (1963). *Science, perception and reality*. Atascadero, CA: Ridgeview Publishing Company.

- Sosa, E. (1991). *Knowledge in Perspective*. Cambridge: Cambridge University Press.
- Sosa, E. (1999). How to defeat opposition to Moore. *Philosophical Perspectives*, 13, 141–153.
- Sosa, E. (2007). *A virtue epistemology: Apt belief and reflective knowledge, volume I*. Oxford University Press.
- Sripada, C. S., & Stanley, J. (2012). Empirical tests of interest-relative invariantism. *Episteme*, 9(01), 3–26. doi:10.1017/epi.2011.2
- Stanley, J. (2005). *Knowledge and Practical Interests*. Oxford University Press.
- Starmans, C., & Friedman, O. (2012). The folk conception of knowledge. *Cognition*, 124(3), 272–283. doi:10.1016/j.cognition.2012.05.017
- Stewart, C. (1988). How to be a Fallibilist. *Philosophical Perspectives*, 2, 91–123.
- Stine, G. C. (1976). Skepticism, relevant alternatives, and deductive closure. *Philosophical Studies*, 29, 249–261.
- Stroud, B. (1984). *The significance of philosophical skepticism*. Oxford: Clarendon Press.
- Swain, M. (1981). *Reasons and knowledge*. Ithaca, NY: Cornell University Press.
- Swain, S., Alexander, J., & Weinberg, J. M. (2008). The Instability of Philosophical Intuitions: Running Hot and Cold on Truetemp. *Philosophy and Phenomenological Research*, 84(1), 138–155.
- Turri, J. (2011). The Express Knowledge Account of Assertion. *Australasian Journal of Philosophy*, 89(1), 37–45. doi:10.1080/00048401003660333
- Turri, J. (2012). Review of: *Achieving Knowledge: A virtue-theoretic account of epistemic normativity*, by John Greco. *Mind*, 121(481), 183–187. <http://doi.org/10.1093/mind/fzs032>

- Turri, J. (2013). A conspicuous art: putting Gettier to the test. *Philosophers' Imprint*, 13(10), 1–16.
- Turri, J. (2014). The problem of ESEE knowledge. *Ergo*, 1(4), 101–127.
- Turri, J. (2015a). Unreliable knowledge. *Philosophy and Phenomenological Research*, 90(3), 529–545. doi:10.1111/phpr.12064
- Turri, J. (2015b). From virtue epistemology to abilism: theoretical and empirical developments. In C. B. Miller, M. R. Furr, A. Knobel, & W. Fleeson (Eds.), *Character: new directions from philosophy, psychology, and theology* (pp. 315–330). Oxford University Press.
- Turri, J. (2015c). An open and shut case: epistemic closure in the manifest image. *Philosophers' Imprint*, 15(2), 1–18.
- Turri, J. (2015d). Skeptical appeal: the source-content bias. *Cognitive Science*, 39(2), 307–324. <http://doi.org/10.1111/cogs.12153>
- Turri, J. (in press a). Primate social cognition and the core human knowledge concept. In E. McCready, M. Mizumoto, J. Stanley, & S. Stich (Eds.), *Epistemology for the rest of the world: linguistic and cultural diversity and epistemology*. Oxford University Press.
- Turri, J. (in press b). Knowledge and assertion in “Gettier” cases. *Philosophical Psychology*.
- Turri, J. (in press c). Vision, knowledge, and assertion. *Consciousness and cognition*.
- Turri, J., & Friedman, O. (2014). Winners and losers in the folk epistemology of lotteries. In J. R. Beebe (Ed.), *Advances in experimental epistemology*. Continuum.
- Turri, J., Buckwalter, W., & Blouw, P. (2014). Knowledge and luck. *Psychonomic Bulletin & Review*. doi:10.3758/s13423-014-0683-5

- Unger, P. (1975). *Ignorance: a case for skepticism*. Oxford: Oxford University Press.
- Unger, P. (1984). *Philosophical relativity*. Oxford: Oxford University Press.
- Vogel, J. (1990). Are there counterexamples to the closure principle? In M. D. Roth & G. Ross (Eds.), *Philosophical Studies Series* (Vol. 48, pp. 13–27). Springer. doi: 10.1007/978-94-009-1942-6\_2
- Vogel, J. (2000). Reliabilism leveled. *The Journal of Philosophy*, 97(11), 602–623.
- Weinberg, J. M., Nichols, S., & Stich, S. (2001). Normativity and epistemic intuitions. *Philosophical Topics*, 29(1&2), 429–460.
- Williamson, T. (2000). *Knowledge and its limits*. Oxford: Oxford University Press.
- Zagzebski, L. T. (1996). *Virtues of the mind: an inquiry into the nature of virtue and the ethical foundations of knowledge*. Cambridge: Cambridge University Press.
- Zamir, E., Ritov, I., & Teichman, D. (in press). Seeing is believing: the anti-inference bias. *Indiana Law Journal*. doi:10.2139/ssrn.1989561