


**Knowledge attributions in iterated fake barn cases**

**JOHN TURRI**

1. **Introduction**

In recent decades, philosophers have frequently discussed ‘fake barn’ cases. (Goldman 1976 introduced the original case into the literature; it is credited to Carl Ginet.) In a fake barn case, the agent believes that something is true because she directly perceives it. But it turns out that she is in an environment where her perceptual evidence could very easily have been misleading and led her to form a false belief. For example, suppose Sarah looks out her car window and sees a roadside barn as she drives along. Naturally she thinks that it is a barn. Everything about Sarah and the barn is normal. But Sarah does not realize that the area she is driving through is being used as a movie set and the set designers have constructed many façades that look just like...
real barns. Sarah is looking at the one real barn among many nearby fakes. Intuitively, many philosophers claim, in such an environment, Sarah does not know that it is barn (e.g. Goldman 1976; Kvanvig 2008: 274; Neta and Rohrbaugh 2004: 401; Pritchard 2005: 161–2; Sosa 1991: 238–9).

Philosophers have relied on this verdict in order to evaluate or motivate theories of knowledge. For example, some argue that it refutes a simple ‘justified true belief’ account of knowledge (Chisholm 1989: 93). The ‘relevant alternatives’ theory of knowledge was motivated by the claim that people are “strongly inclined” to deny knowledge in fake barn cases (Goldman 1976: 772 ff.). Similarly, an ‘anti-luck’ theory of knowledge has been defended on the grounds that it explains the intuition that knowledge is absent in such cases (Pritchard 2012). Some philosophers have rejected this intuition or even offered arguments that knowledge is present (e.g. Hetherington 1999; Lycan 2006; Millikan 1984; Turri 2012). There might even be evidence, some have argued, that philosophers implicitly attribute knowledge in fake barn cases but have been trained to reverse this judgment explicitly (Turri 2016a). Nevertheless, textbooks and review articles treat it as uncontroversial that knowledge is absent in fake barn cases (Chisholm 1989; Engel 2015: Ichikawa & Steup 2012; Luper 2016). A theory that implies otherwise is considered ‘revisionary’ (Jarvis 2013: 531).

Researchers have recently investigated the knowledge judgments that people actually make about fake barn cases. The results have consistently shown that people tend to attribute knowledge to agents in such cases. For example, one research team tested a case involving an agent who sees a real house amidst a large number of ‘house façades’ (Colaco et al. 2014). People tended to judge that the agent knows he sees a house. Another research team tested cases with a fake barn structure that involved the detection of rodents or gems. People tended to attribute knowledge of the rodent or gem, whereas people in closely matched controls denied knowledge when the agent did not detect the relevant item (Turri et al. 2015). In another series of studies, people tended to attribute knowledge that an albino vervet monkey is in a tree even when it is ‘surrounded by’ visibly indistinguishable snow monkeys (Turri 2016b). Researchers have even tested cases involving “fake barn façades” very closely modeled after the original one from the epistemological literature. People overwhelmingly judged that the agent knows it is a barn, and they did so at rates comparable to those observed in a closely matched control condition involving ‘cheap’ barns, which did not even mention fakes (Turri 2016c). Ironically, researchers have even found that most professional philosophers attribute knowledge in cases with a fake barn structure (Horvath and Wiegmann 2016).

Thus, there is a substantial body of research establishing that most people, including philosophers, naturally attribute knowledge in fake barn cases. However, with one exception to be noted shortly, researchers have tested what might be called single-iteration fake barn cases. In a single-iteration case, the agent does not encounter any of the nearby fakes and, thus, has
not been mislead by any of them. Instead, she is described as encountering just one item – the real item of interest – which means that she forms a true belief, is never mislead, and forms no false beliefs. By contrast, in a *multiple-iteration* case, the agent first encounters several nearby fakes, is mislead by them, forms a false belief each time, and then encounters and correctly identifies the real item of interest. According to some theorists, single-iteration cases do not effectively elicit the judgment that knowledge is absent. Instead, they claim, multiple-iteration cases should be used because they naturally lead us to judge ‘quite strongly’ that knowledge is absent (DeRose 2009: 23, n. 24).

One study partially addressed this. Researchers manipulated whether participants were told that the agent ‘has not yet encountered any’ of the nearby fakes, or ‘has already encountered a large number’ of them (Colaço et al. 2014). This manipulation did not affect knowledge attributions and participants tended to attribute knowledge in both cases. However, the manipulation does not fully capture what is supposed to be important about the multiple-iteration case. In particular, it is not made clear that the encountered fakes actually mislead the agent and caused her to form false beliefs.

The present study was designed to investigate knowledge judgments in a more fully described multiple-iteration fake barn case. In the more fully described case, the agent encounters four fakes and verbally misidentifies each as an item of interest, after which she encounters and correctly identifies a genuine item of interest.

To provide a fuller context for interpreting the results from the multiple-iteration fake barn case, I also did three other things. First, I compared a multiple-iteration fake barn case to a single-iteration fake barn case. I accomplished this by manipulating (between-subjects) whether the agent correctly identified an item of interest on the first try or on the fifth try (after four explicit failures). The purpose of this is to determine whether iterating errors has the alleged effect on people’s knowledge judgments in fake barn cases.

Second, I compared cases involving nearby fakes to cases without any fakes. I accomplished this by manipulating (between-subjects) whether nearby fakes are mentioned or not. The purpose of this is to determine whether iterating error affects knowledge judgments differently depending on whether there are nearby fakes, or if its effect is consistent.

Third, I collected four other judgments (within-subjects) that many researchers assume are directly relevant to knowledge attributions. In particular, participants also rated whether the target proposition was true (truth evaluation), whether the agent thinks it is true (belief attribution), whether the agent has good evidence for it (evidence evaluation), and whether the agent is able to detect items of the relevant kind (ability attribution). The

---

1 See also the comments on this weblog post: http://certaindoubts.com/what-philosophers-think-might-not-be-what-you-think-they-think/
purpose of this is to gain insight into the underlying psychological processes informing knowledge judgments. Collecting these judgments provides evidence about whether knowledge judgments are affected because these other judgments are affected. In other words, it can help us determine whether the other judgments mediate an observed effect on knowledge judgments.

2. Experiment

2.1 Method

Participants\(^2\) were randomly assigned to one of four conditions in a 2 (presence: unspecified, specified) × 2 (iteration: single, multiple) between-subjects design. All participants read a simple story about an agent, Sarah, who is asked about barns while driving through the countryside (adapted from Turri 2016c). The presence factor varied whether the presence of nearby fakes was not mentioned at all (unspecified) or explicitly mentioned (specified). The iteration factor manipulated whether the agent correctly identified an object of interest, a barn, on the first try (single), or misidentified an object of interest four times before correctly identified one the fifth time (multiple). Here is the first paragraph of the story (the unspecified/specifed manipulation is bracketed):

Sarah is driving with her son down the highway. Sarah looks out the window of her car and sees many structures. Sarah doesn’t realize that the countryside she is driving through [was previously used as the set of a film, and that many famous movie directors have considered it to be an ideal place to shoot movies that involve rural action and romance/is currently being used as the set of a film, and that the set designers have constructed many fake barn facades in this area that look as though they are real barns]. Sarah’s son says “Mom, I have to do a report on barns for my social studies class. Could you show me a barn?”

In the single conditions, the story then ended with a short second paragraph (the unspecified/specifed difference is bracketed):

The first structure Sarah sees looks like a barn. Despite all the [structures/fakes] around, Sarah is in fact looking at the one real barn in the area. She says, “That’s a barn.”

In the multiple conditions, the story instead ended with five short paragraphs (the unspecified/specifed difference is bracketed):

The first structure Sarah sees looks like a barn but it isn’t. She says, “That’s a barn.”

\(^2\) \(N = 163; \) aged 18–74; mean age = 35 years; 53 female; 94% reporting English as a native language. Participants were US residents, recruited and tested online using Amazon Mechanical Turk (AMT) and Qualtrics, and compensated $0.40 for approximately 2 min of their time. Repeat participation was prevented (by AMT worker ID).
The second structure Sarah sees looks like a barn but it isn’t. She says, “That’s a barn.”

The third structure Sarah sees looks like a barn but it isn’t. She says, “That’s a barn.”

The fourth structure Sarah sees looks like a barn but it isn’t. She says, “That’s a barn.”

The fifth structure Sarah sees looks like a barn. Despite all the [structures/fakes] around, Sarah is in fact looking at the one real barn in the area. She says, “That’s a barn.”

After reading the story, participants rated five test statements in a matrix table (order rotated randomly):

- It is true that it’s a barn. (truth)
- Sarah thinks that it’s a barn. (thinks)
- Sarah has good evidence that it’s a barn. (evidence)
- Sarah is able to detect barns. (ability)
- Sarah knows that it’s a barn. (knows)

Responses were collected on a standard 7-point Likert scale, 1 (‘strongly disagree’)–7 (‘strongly agree’), left-to-right on the participant’s screen. The story remained atop the screen throughout testing. After testing, participants completed a short demographic survey.

2.2 Results

Table 1 includes descriptive statistics for all the dependent variables. An analysis of variance on knowledge judgments revealed no main effect of presence (i.e. whether the presence of fakes was unspecified or specified), a main effect of iteration (i.e. whether the agent made no mistakes or four mistakes before correctly identifying the object of interest), and an interaction of presence and iteration (see Appendix, Table A2). The interaction indicates that the effect of iteration on knowledge judgments differed depending on whether the presence of fakes was unspecified or specified. To explore this interaction, I conducted follow-up independent samples t-tests. In unspecified conditions, mean knowledge judgment was significantly higher in the single condition than in the multiple condition and the mean difference was very large\(^3\) (see Figure 1). By contrast, in specified conditions, mean knowledge judgment did not differ between the single condition and multiple condition.\(^4\)

One sample t-tests revealed that mean knowledge judgment was trending below the neutral midpoint in the unspecified multiple condition, but it

---

\(^3\) \(t(79) = 5.74, p < .001\) (all reported tests two-tailed), \(\text{MD} = 1.94\ [1.27, 2.61], d = 1.29\).

\(^4\) \(t(80) = 0.65, p = .515\) ns.
was significantly above the midpoint in the other three conditions (see Figure 1 and Appendix, Table A3).

In order to gain insight into the psychological processes informing knowledge judgments, I conducted multiple linear regression analyses. Because of the interaction effect on knowledge judgments, I conducted separate regression analyses predicting knowledge judgments in the unspecified and specified conditions. For each analysis, the predictors were assignment to iteration condition (single, multiple), truth evaluation, belief attribution, evidence evaluation, ability attribution, participant gender and participant age. In each case, the model explained most of the variance in knowledge judgments ($R^2$ values were .611 and .548) (see Appendix, Table A4.) In the unspecified conditions, only evidence evaluations and ability attributions predicted knowledge judgments. These predictive relationships were positive (e.g. higher ability attributions predicted higher knowledge judgments, and lower ability attributions predicted lower knowledge judgments). In the specified conditions, evidence evaluations, ability attributions and truth evaluations predicted knowledge judgments. Again, the predictive relationships were positive. In neither model did participant age, participant gender, belief attributions, or assignment to iteration condition make a unique significant contribution to predicting knowledge judgments.

As noted above, analysis of variance indicated that assignment to iteration condition affected knowledge judgments. However, regression analysis indicated that assignment to iteration condition did not make a unique predictive contribution to knowledge judgments. That is, when controlling for the effect of the other four judgments, iteration had no additional effect. This suggests that the effect of iteration on knowledge judgments is mediated by its effect on the other

Figure 1. Mean knowledge judgments in the four conditions. Scales ran 1 (SD) – 7 (SA). Error bars represent bootstrapped 95% confidence intervals.
judgments collected. To directly confirm this, I conducted bootstrap mediation analyses (Hayes 2013) on knowledge judgments in the unspecified conditions, which is where iteration had its effect (as indicated by the independent samples t-tests reported above). Evidence evaluations and ability attributions were the only significant predictors of knowledge judgments in the unspecified conditions, so I conducted a separate mediation analysis for each (see Figure 2).

For the first analysis, I entered assignment to iteration condition as the independent variable, knowledge judgment as the outcome, and evidence evaluation as potential mediator. Evidence evaluations mediated the effect of condition. The ratio of indirect effect to direct effect was 1.01 and the mediation’s effect size was large. For the second analysis, I did the exact same thing except that the potential mediator was ability attribution. Ability attributions completely mediated the effect of condition. The ratio of indirect effect to direct effect was 5.16 and the mediation’s effect size was very large.

The mediation analyses just reported assume a causal direction. In particular, they assume that knowledge judgments were based on evidence evaluations and ability attributions. However, in each case the reverse mediation model was also significant. It is thus possible that evidence evaluations and ability attributions were actually based on knowledge judgments, rather than vice versa. In order to help decide which of these possibilities better fits the data, I conducted two causal searches with the Greedy Equivalence Search (GES) algorithm. Similar to regression-based mediation analysis, GES is used

Figure 2. Mediation analyses examining the effect of iteration (single, multiple) on knowledge judgments via evidence evaluations (Panel A) and ability attributions (Panel B) in the unspecified conditions. Reference class for iteration: single. Parenthetical values represent the strength of a simple regression between the two variables; values outside the parentheses represent the strength of the relationships in a model used to test for mediation. **p ≤ .01, ***p ≤ .001.

Indirect effect = -0.97 [-1.56, -0.51]; direct effect = 0.97 [-1.58, -0.34]; $\kappa^2 = .29 [ .15, .44]$. The $\kappa^2$ (kappa-squared) index for mediation analyses can range between 0 and 1 and represents ‘the proportion of the maximum possible indirect effect that could have occurred’, given the experimental design and data (Preacher and Kelley 2011). I follow previous researchers in defining small, medium, and large effect sizes for $\kappa^2$ as .01, .09 and .25 (Preacher and Kelley 2011).

Indirect effect = -1.62 [-2.39, -1.00]; direct effect = -0.31 [-1.12, 0.49]; $\kappa^2 = .39 [ .24, .55]$. 

5 Indirect effect = -0.97 [-1.56, -0.51]; direct effect = 0.97 [-1.58, -0.34]; $\kappa^2 = .29 [ .15, .44]$. The $\kappa^2$ (kappa-squared) index for mediation analyses can range between 0 and 1 and represents ‘the proportion of the maximum possible indirect effect that could have occurred’, given the experimental design and data (Preacher and Kelley 2011). I follow previous researchers in defining small, medium, and large effect sizes for $\kappa^2$ as .01, .09 and .25 (Preacher and Kelley 2011).

6 Indirect effect = -1.62 [-2.39, -1.00]; direct effect = -0.31 [-1.12, 0.49]; $\kappa^2 = .39 [ .24, .55]$. 


to make causal inferences from correlations and independence relations in a dataset. However, whereas mediation analysis assumes a causal direction, GES does not. Instead, GES considers all possible models available given the different variables and selects the one with the best information score (for more on GES and the general theory behind causal search, see Chickering 2002; Meek 1997; Pearl 2000; Spirtes et al. 2000. For prior applications of GES in philosophical debates, see Rose and Nichols 2013; Rose et al. 2012; Turri 2016b; Turri and Turri in press).

The causal searches were conducted with Tetrad 5 (http://www.phil.cmu.edu/projects/tetrad/). For the first search, I entered assignment to iteration condition, evidence evaluation, and knowledge judgment into GES. The model was constrained so that assignment to condition could not be caused by either dependent variable. Figure 3A depicts the best fitting model, which fit the data well.\(^7\) In the model, knowledge judgments cause evidence evaluations. For the second search, I did exactly the same thing except for replacing evidence evaluations with ability attributions. Figure 3B depicts the best fitting model, which also fit the data well.\(^8\) In the model, ability attributions cause knowledge judgments.

3. Conclusion

The purpose of this experiment was to investigate knowledge judgments in multiple-iteration fake barn cases. In a single-iteration fake barn case, the agent correctly identifies an object of interest on the first try, despite the presence of nearby fakes that easily could have mislead her. In a multiple-iteration fake barn case, the agent first encounters several fakes, misidentifies each of them and then encounters and correctly identifies a genuine item of

\[\chi^2(1) = 2.36, \ p = .125, \ \text{BIC} = -2.03.\]

\[\chi^2(1) = 0.62, \ p = .431, \ \text{BIC} = -3.77.\]
interest. Prior work has established that people tend to attribute knowledge in single-iteration cases, but multiple-iteration cases have not been tested.

The central finding was that people tend to attribute knowledge in multiple-iteration fake barn cases. Moreover, although mean knowledge attribution was numerically lower in multiple-iteration fake barn cases than in the single-iteration cases, this difference was not statistically significant. By contrast, the iteration of errors had a large effect on knowledge attributions in a pair of closely matched control cases that did not involve fakes. In other words, the presence of nearby fakes could prevent iterated errors from affecting people’s knowledge judgments. The results from mediation and causal search analyses suggest an explanation for this. When the agent repeatedly misidentifies structures as barns and there is no external explanation for her errors, people conclude that she is unable to identify barns, which leads them to doubt that she knows. By contrast, when the agent repeatedly misidentifies structures as barns, the presence of fakes is an obvious external explanation for her errors, so people do not conclude that she is unable to identify barns. This coheres with recent findings showing that when someone forms a true belief, people’s default assumption is that she formed the true belief through a relevant cognitive ability and thus knows (Turri 2016d).

If the preceding explanation is on the right track, then it suggests a potentially fruitful avenue of research moving forward. If knowledge judgments are guided by ability attributions, then some disagreements about knowledge will be due to underlying disagreements about abilities. This could explain some disagreements among philosophers about knowledge in fake barn cases. For example, some philosophers claim that knowledge is (roughly) true belief formed through cognitive ability, and they attribute knowledge in fake barn cases (e.g. Turri 2011; Turri 2015a). Other philosophers also agree that knowledge is true belief formed through cognitive ability, but they deny knowledge in fake barn cases (e.g. Greco 2010: Chapter 5; Pritchard 2012). And it just so happens that these philosophers who disagree about knowledge in fake barn cases also disagree about whether the agent has an ability to detect barns. Interestingly, the latter disagreement can be traced to a very specific question about the metaphysics of abilities, namely, whether abilities must be reliable, in the sense that they must produce mostly true beliefs (see Turri 2015b; Turri in press). Those who deny knowledge in fake barn cases claim that abilities require reliability, whereas those who attribute knowledge claim that abilities can be unreliable.

Thus, one question for future research is whether – and if so, how consistently – disagreements in epistemological judgments can be traced to underlying differences in metaphysical judgments. Perhaps there are important individual differences in how much unreliability is viewed as consistent with having an ability and, thus, with gaining knowledge. Future work could investigate what these differences are and whether philosophers tend to have traits that predict greater intolerance of unreliability (compare Feltz and
Cokely 2012). The present results suggest that neither gender nor age contributes to this intolerance in a consistent or meaningful way.9

Funding

This work was supported by the Social Sciences and Humanities Research Council of Canada; the Ontario Ministry of Economic Development and Innovation; and the Canada Research Chairs program.

Department of Philosophy and Cognitive Science Program
University of Waterloo
200 University Avenue West
Waterloo, Ontario N2L3G1, Canada
john.turri@gmail.com

References


9 For helpful feedback and discussion, I thank David Colaço, Jeremy Fantl, Joachim Horvath, and Angelo Turri.


**Appendix**

**Table A1.** Descriptive statistics for all dependent variables in the four conditions

<table>
<thead>
<tr>
<th>Measure</th>
<th>Unspecified</th>
<th>Specified</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Single</td>
<td>Multiple</td>
</tr>
<tr>
<td></td>
<td>(N = 41)</td>
<td>(N = 40)</td>
</tr>
<tr>
<td>Truth</td>
<td>Mo 7</td>
<td>Md 7</td>
</tr>
<tr>
<td>Thinks</td>
<td>Mo 7</td>
<td>Md 7</td>
</tr>
<tr>
<td>Evidence</td>
<td>Mo 6</td>
<td>Md 6</td>
</tr>
<tr>
<td>Ability</td>
<td>Mo 6</td>
<td>Md 6</td>
</tr>
<tr>
<td>Knows</td>
<td>Mo 6</td>
<td>Md 6</td>
</tr>
</tbody>
</table>

N, number of participants; Mo, mode; Md, median; M, mean; SD, standard deviation.

**Table A2.** Univariate analysis of variance of all dependent variables

<table>
<thead>
<tr>
<th>Measure</th>
<th>Presence</th>
<th>Iteration</th>
<th>Presence x iteration</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>F df p</td>
<td>F df p</td>
<td>F df p</td>
</tr>
<tr>
<td>Truth</td>
<td>0.36 1, 159 .551 .002</td>
<td>3.80 1, 159 .053 .23</td>
<td>0.30 1, 159 .586 .002</td>
</tr>
<tr>
<td>Thinks</td>
<td>0.25 1, 159 .617 .002</td>
<td>0.01 1, 159 .981 .001</td>
<td>3.65 1, 159 .058 .022</td>
</tr>
<tr>
<td>Evidence</td>
<td>4.91 1, 159 .028 .030</td>
<td>16.45 1, 159 &lt;.001 .094</td>
<td>9.32 1, 159 .003 .055</td>
</tr>
<tr>
<td>Ability</td>
<td>3.99 1, 159 .047 .025</td>
<td>49.92 1, 159 &lt;.001 .239</td>
<td>12.79 1, 159 &lt;.001 .074</td>
</tr>
<tr>
<td>Knows</td>
<td>1.86 1, 159 .175 .012</td>
<td>17.96 1, 159 &lt;.001 .102</td>
<td>10.55 1, 159 .001 .062</td>
</tr>
</tbody>
</table>

**Table A3.** One sample t-tests on mean knowledge judgments in the four conditions (test value = 4)

<table>
<thead>
<tr>
<th>Iteration</th>
<th>Unspecified</th>
<th>Specified</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>t df p</td>
<td>MD 95% CI</td>
</tr>
<tr>
<td>Single</td>
<td>6.99 40 &lt;.001 1.49 1.06, 1.92 1.10</td>
<td>4.53 42 &lt;.001 1.00 0.55, 1.45 0.69</td>
</tr>
<tr>
<td>Multiple</td>
<td>-1.71 39 .095 -0.45 -0.98, 0.08 0.27</td>
<td>2.24 38 .031 0.74 0.07, 1.42 0.36</td>
</tr>
</tbody>
</table>