the british psychological society
promoting excellence in psychology

British Journal of Social Psychology (2021), 60, 808–825

www.wileyonlinelibrary.com

Men have ability, women are lucky: A pre-registered experiment examining gender bias in knowledge attribution

Natalie G. Disher, Antonio L. Guerra and Gerald J. Haeffel
Department of Psychology, The University of Notre Dame, Indiana, USA

The purpose of this experiment was to test the effect of gender on knowledge attribution using a Justified True Belief (JTB) framework. A 2 (gender: male, female) × 3 (knowledge case: knowledge control, Gettier, ignorance control) experimental design was used with a sample of 420 U.S. adults. Contrary to hypotheses, participants attributed similar levels of knowledge to male and female agents across all knowledge conditions; participants also rated males and females as equally likely to have the 'right' answer across knowledge conditions. However, knowledge was more likely to be attributed to luck (as opposed to ability) for female agents than it was for male agents across knowledge conditions and scenarios. This result suggests that while overt forms of gender bias may be fading, more covert forms still exist. Secondary analyses also showed that the JTB methodology was not robust to scenario type or knowledge condition. Comprehension was affected by both knowledge condition (knowledge vs. Gettier/ignorance) and the specific content of the hypothetical scenario presented (e.g., squirrel vs. jewelery). These confounds should be addressed in future JTB studies as it is possible that differences between knowledge and Gettier cases are due to misunderstanding the scenarios rather than beliefs about knowledge. Theoretical and practical implications of the findings are discussed.

As humans, we often have to make judgements about what other people know. These judgements are called knowledge attributions (Turri, Buckwalter, & Blouw, 2015), and they can affect our attitudes, perceptions, and even the opportunities we afford other people (Steinpreis, Anders, & Ritzke, 1999). According to the Justified True Belief theory (JTB), there are three requirements to justify an individual's knowledge claim (Gettier, 1963). The individual must believe a fact to be true, the fact must indeed be true, and the individual must be justified in his or her belief. These conditions, first proposed by Plato, have gained traction with both philosophers and members of the psychology community who study knowledge and knowledge attribution (Nagel, Juan, & Mar, 2013).

Gettier (1963), however, identified unique cases in which a person may meet the JTB requirements, yet is not viewed by others as possessing knowledge. In these Gettier cases, the individual makes a correct factual conclusion, but the circumstance for why the conclusion is correct is different from what the person believes. For example, imagine that a woman arrives home from work and sees her husband sitting in the living room. What she is unaware of, however, is that the man she sees is actually her husband's twin brother.

*Correspondence should be addressed to Gerald J. Haeffel, University of Notre Dame, Corbett Hall, Notre Dame, IN 46656, USA (email: ghaeffel@nd.edu).
Pre-registration (anonymous for peer-review): https://osf.io/xtaj4/?view_only=1aae719c8ef94d91b0d69fe8fc8f7c63

DOI:10.1111/bjso.12443

Her husband is also sitting in the living room, but in the adjacent chair. She is correct that her husband is in the living room, but she is incorrect in her original assumption: that her husband is in the chair that the twin brother is in (Turri, 2016; Turri et al., 2015). In these cases, the agents make correct assertions, but these assertions may be considered lucky rather than true knowledge.

Research shows that, despite meeting the JTB requirements for possessing knowledge, Gettier cases tend to be viewed as distinct from true knowledge cases. Studies (e.g., Nagel et al., 2013) show that knowledge in Gettier cases is attributed more to luck than true knowledge (although some studies have found no differences between Gettier and true knowledge cases; Turri et al., 2015). A limitation of this research, however, is that it has failed to account for possible pre-existing beliefs that can bias knowledge attribution. For example, factors such as sexism may affect how people view the knowledge of others.

Research shows that the judgements people make about others' competence and expertise is affected by whether or not the target is a male or female (Alexander & Andersen, 1993; Bigelow, Lundmark, Parks, & Wuebker, 2012; Bosak & Sczesny, 2011; Steinpreis et al., 1999). For example, simply changing the gender of the name on a resume or curriculum vitae can affect how favourably potential employers view the candidate (Bosak & Sczesny, 2011; Moss-Racusin, Dovidio, Brescoll, Graham, & Handelsman, 2012; Steinpreis et al., 1999). Resumes with a female name are viewed less positively than resumes with a male name (despite identical qualifications). Similarly, studies have demonstrated bias against female political candidates as well as professors (Alexander & Andersen, 1993; Basow, 1995). For example, Basow (1995) found that for male professors, the gender of the evaluator did not affect the results of the evaluation, but for female professors, the most favourable evaluations came from female students whereas the least favourable evaluations came from male students (Basow, 1995).

Research also shows that there are differences in how others view success for males and females. For example, Espinoza, Da Luz, Fontes, and Arms-Chavez (2014) found that men's successes in math were more likely to be attributed to ability whereas women's successes in math were more likely to be attributed to effort. Similarly, Proudfoot, Kay, and Koval (2015) found that ideas were more likely to be viewed as ingenious if they were expressed by a man rather than by a woman. This study also found that men were thought to be more creative than women and were more capable of thinking outside the box. These findings may help explain why women are not promoted to leadership positions even when they have demonstrated a greater level of competency than their male counterparts (Heilman, 2001).

Taken together, this work suggests that gender plays an important role in how others view knowledge, competence, and ability. Thus, it is important to consider this factor in studies of knowledge attribution. Fortunately, the JTB and Gettier case framework is perfectly suited for studying such biases. JTB theory and Gettier cases provide a novel strategy for investigating a continuum of gender bias ranging from overt to more covert. The true knowledge scenarios used in JTB research are unequivocal in their description of what knowledge the target possesses and the accuracy of that knowledge (i.e., the three JTB requirements). In these cases, there is a good deal of contextual information that constrains knowledge attributions. This means that a high level of bias might be required to directly contradict such information. In contrast, Gettier cases are more ambiguous and thus may allow for the detection of more subtle biases. Further, Gettier cases may have greater external validity than the true knowledge cases because the information available to people in real life tends to be limited. People do not have access to the inner cognitive processes of others, which makes it difficult to determine if they are justified in their

beliefs. To this end, many real-life situations might be evaluated as Gettier-like situations. For example, imagine a male and female student both got a difficult problem correct on a multiple-choice math test. In this situation, both students believed that they had the correct answer, they were justified in that belief because they had studied, and their answer was, indeed, correct. However, the teacher would not necessarily know how much the students studied or their certainty in their answer. Thus, the teacher may still infer that the female student arrived at the correct answer through faulty logic or guesswork while the male student arrived at the answer because of his strong mathematical abilities.

Identifying the extent to which knowledge attribution is affected by gender bias is important because these biases have the potential to affect people's self-efficacy, pursuit of academic degrees in STEM, and occupational success. Further, testing the effect of gender on knowledge attributions may have methodological implications for future research on JTB theory. Prior research has generally supported the distinctiveness of the Gettier case, but there are also failed replications (e.g., Turri et al., 2015). One explanation for the inconsistencies in prior research is that studies have used different knowledge scenarios (Hall et al., 2018). These scenarios differ in regard to the gender of the agent character as well as in content (e.g., squirrels, houses, jewelery). Researchers have assumed that knowledge attributions are robust to these methodological differences, but this assumption has yet to be tested. It is possible that differences in scenario content (e.g., the proportion of male and female agents, type of story) have led to the inconsistent findings in the JTB literature.

The purpose of this study was to fill this gap in the literature and test the effect of gender on knowledge attributions. Participants were randomized to one of six conditions using a 3 (knowledge condition: knowledge/ignorance/Gettier) $\times$ 2 (gender condition: male name/female name) between-participants experimental design. The outcomes of interest were measured with three knowledge probes used in prior JTB studies. The first probe assessed the degree to which the participant judged the target to have knowledge. The second probe assessed whether the participant viewed the target to be right or wrong. The third probe assessed the degree to which the target's knowledge was due to luck or ability. Participants also completed a comprehension probe, which served as a manipulation check to ensure that they were reading the scenarios.

We had three *primary hypotheses*: (1) Participants would attribute similar amounts of knowledge to male and female agents in the true knowledge and ignorance conditions. However, in the more ambiguous Gettier condition, participants would attribute less knowledge to agents with female names than to agents with male names. (2) Participants would be equally likely to judge the male and female targets as having the right answer in the true knowledge and ignorance conditions. However, in the more ambiguous Gettier condition, participants would be more likely to judge the conclusion of agents with female names as being wrong than those of agents with male names. (3) Participants would attribute similar amounts of luck/ability in the true knowledge and ignorance conditions. However, in the more ambiguous Gettier condition, participants would be more likely to attribute knowledge to luck for agents with female names than with male names.

We also conducted a set of *secondary analyses* to test the effect of scenario type on knowledge attributions. As noted earlier, some studies have failed to find a distinction between true knowledge and Gettier cases. One possible explanation for the inconsistencies is that researchers have not used a standard set of knowledge scenarios; it is possible that different scenarios elicit different knowledge attributions from participants. However, researchers have yet to specify what content (or how that content) would lead

to different attributions. Thus, we hypothesized that scenario type would not affect knowledge attributions.

Method

Open practices

Pre-registration (anonymous for peer-review) can be found here: https://osf.io/xtaj4/?view_only=1aae719c8ef94d91b0d69fe8fc8f7c63. Materials can be found here: https://osf.io/uvb83/. Data can be found here: https://osf.io/b3jfw/. Code is available upon request.

Power analysis

An a priori power analysis (conducted using G*Power 3.1) with six groups showed that a total of 390 participants were needed to detect a small to medium effect size ($f = .2$) with a power of .95 and alpha level $p < .05$. Sample size was determined before any data analysis.

Participants

Participants were 420 adults ages 18 to 77 (mean age = 35; 227 males, 189 females, four unspecified). They were recruited using the Prolific online study platform (Peer, Brandimarte, Samat, & Acquisti, 2017) and completed the experiment online. All participants were from the United States to ensure similar familiarity of gendered names (which were chosen from the United States census). Seventy-four per cent of participants self-identified as White/European descent ($n = 310$), 11.0% Asian descent ($n = 46$), 6.2% Latin/Hispanic descent ($n = 26$), 5.7% Black/African descent ($n = 24$), and less than 2% of participants endorsed Southeast Asian descent, Hawaiian/Pacific Islander descent, and 'other'. All research procedures were fully consistent with APA ethical guidelines, and the study was approved by the University's Human Subjects Committee. Participants were paid an average of $9.52 per hour to complete the 5-min experiment.

Materials

Knowledge conditions

Participants were randomly assigned to one of three knowledge conditions: Knowledge, Ignorance, or Gettier. In each knowledge condition, participants read a two-paragraph scenario (see Appendix for all knowledge condition paragraphs). The first paragraph was the same for all conditions, but each participant read a different paragraph depending on which of the three knowledge conditions to which they were randomly assigned (knowledge, ignorance, or Gettier). In the knowledge condition, the agent in the scenario forms a true belief based on a true state of affairs (in the absence of any threat to disrupt the true scenario); in the ignorance condition, the agent in the scenario forms a false belief based on a successful/valid threat that negates the true scenario; and, in the Gettier condition, the agent in the scenario forms a true belief based on a true state in the presence of a salient but ultimately failed threat of disruption to the scenario.

Participants were randomly assigned to one of three scenario conditions: Squirrel, House, or Jewelery (see Appendix for all scenario paragraphs). In the squirrel scenario, the agent is an ecologist attempting to identify red-speckled ground squirrels in a region in which there are also similar looking animals (prairie dogs). In the house scenario, the

agent is driving with his or her child through a countryside that had been hit by a tornado; the agent is pointing out houses in an area in which there are also realistic-looking fake houses. In the jewelery condition, the agent is shopping for a diamond watch at a jewelery store in which there are also fake diamonds that look identical to real ones. Three scenarios were used to rule out the possible confounds related to the use of a specific scenario; this also allowed us to examine if knowledge attributions were robust to scenario type (tested in the secondary analyses).

Gender name conditions

Participants were randomly assigned to one of two gender name conditions – male name or female name. In the male name condition, the agent in the scenario was randomly called John, Craig, or Brian. In the female name condition, the agent in the scenario was randomly called Melissa, Lori, or Jennifer. The names were chosen from a list generated by Knobloch-Westerwick, Glynn, and Huge (2013) based on the US Social Security Administration's database for popular gender names.

Knowledge probes (dependent variables)

All knowledge probes are listed in the Appendix.

1. *True Knowledge*. Participants were asked to judge on a 100-point sliding scale the degree to which the agent 'Knows' or 'Only Believes' a knowledge statement. For example, in the squirrel scenario, the probe reads: *(Agent name)_____that there is at least one red-speckled ground squirrel in Zone 3 today*; underneath the probe is a visual sliding scale (Only Believes <--------> Knows). The side of the scale in which each descriptor appeared was randomized for each participant. This question was presented in the presence of the knowledge scenario.
2. *Right/Wrong and Luck/Ability*. Participants answered two questions regarding the attribution of luck to the individual. Regardless of the scenario, this question read *(Agent name) got the_____answer because of his/her*. . . . . .The first blank was answered with a multiple-choice question with the options *right/wrong*, and the second blank was answered on a 100-point sliding scale, (In)Ability <----------> Good/Bad Luck. The side of the scale in which each descriptor appeared was randomized for each participant.
3. *Comprehension*. Participants were asked a multiple-choice comprehension question to determine if they read and understood the objective outcome of the story. Consistent with prior research (e.g., Turri et al., 2015), this question was used as a manipulation check. Those who failed to answer the comprehension correctly were excluded from the final analyses. For the squirrel scenario, the probe read: *(Agent name) is looking at a ground squirrel/prairie dog*. This probe was presented on a new page, in the absence of the knowledge scenario.

Procedure

Participants were recruited online via Prolific (Peer et al., 2017). We chose Prolific because samples tend to be more diverse and are more scientifically naive compared to other popular online platforms (Peer et al., 2017). The experiment was created using

Qualtrics. After consenting to participate, participants completed a short demographics questionnaire and were then randomly assigned to one of the six conditions. They were instructed to read the scenario and were then presented with the knowledge probes. It is important to note that participants also completed a probe assessing 'reasonableness' of the target's judgement. We did not have a specific hypothesis for this outcome and, thus, did not include it in the analyses. All other measures, manipulations, and exclusions are reported.

Results

Manipulation check

Sixty-two participants (15%) failed to answer the comprehension probe correctly and were excluded from analyses. This resulted in a total sample size of 358, which was 32 participants fewer than the sample size from the original power analysis. This attrition led to a decrease in the power to detect a small to medium effect from .95 to .93. We conducted *post-hoc* exploratory analyses using binomial logistic regression to determine if there were any systematic patterns to explain differences in comprehension accuracy. There was no significant main effect of age, gender, or ethnicity of the participant on comprehension accuracy. However, there was a significant main effect of gender condition, $\chi^2 = 4.89$, $p = .03$, OR = 1.85. Participants were less likely to answer the comprehension probe correctly when the gender of the agent in the scenario was female (81% correct comprehension) than for when it was male (89% correct comprehension). There was also a main effect of knowledge condition ($\chi^2 = 13.10, p = .001$). Participants were significantly less likely to answer the comprehension probe correctly in the Gettier (82% correct comprehension, $p = .003$, OR = 3.29) and ignorance (81% correct comprehension, $p = .002$, OR = 3.51) conditions than in the knowledge (94% correct comprehension) condition. Finally, there was a significant knowledge condition by scenario interaction ($\chi^2 = 69.1, p < .001$). Participants randomly assigned to the squirrel-ignorance condition accounted for a disproportionately high percentage of the incorrect comprehension responses (34%; $n = 21$).

Hypothesis 1: Knowledge

We hypothesized that participants would attribute similar amounts of knowledge to male and female agents in the true knowledge and ignorance conditions. However, in the more ambiguous Gettier condition, participants would attribute less knowledge to agents with female names than to agents with male names. To test this hypothesis, we used an Analysis of Variance (ANOVA) with knowledge condition and gender condition as the independent variables. The dependent variable was participants' score on the 'knows' → 'only believes' 100-point sliding scale. Higher scores indicated greater knowledge attribution.

Contrary to hypotheses, there was neither a significant main effect ($F[1, 352] = .06$, $p = .80$, $\eta^2 = .00$) nor a moderating effect ($F[1, 352] = .01, p = .99, \eta^2 = .00$) of gender condition. However, there was a significant main effect of knowledge condition, $F[2, 352] = 98.15$, $p < .001$, $\eta^2 = .36$. Replicating prior research, our results supported the distinctiveness of the Gettier condition (see Figure 1). Tukey *post-hoc* comparisons showed that participants in the knowledge condition ($M = 81$) were more likely to attribute knowledge to the agent than participants in the Gettier ($M = 56$; $t = 5.8$, $p$tukey $< .001$) and ignorance conditions ($M = 20$; $t = 13.99$, $p$tukey $< .001$). And,

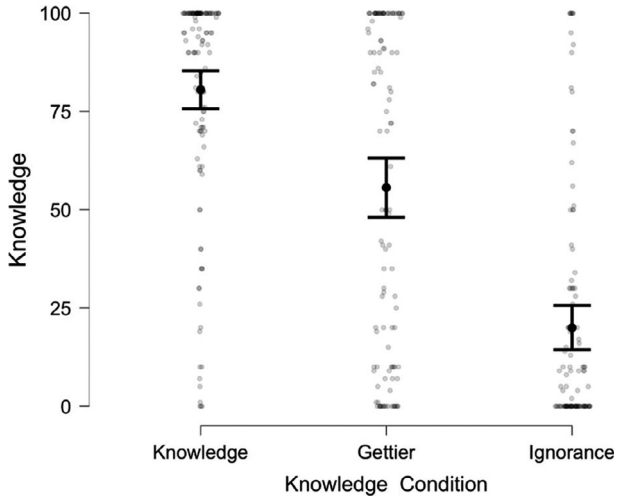

Figure 1. Knowledge rating (0-100) as a function of knowledge condition (with standard error bars).

participants in the Gettier condition were significantly more likely to attribute knowledge to the agent than participants in the ignorance condition (t = 8.0, $p$tukey < .001).

Hypothesis 2: Right/Wrong

We hypothesized that participants would be equally likely to judge the male and female targets as having the right answer in the true knowledge and ignorance conditions. However, in the more ambiguous Gettier condition, participants would be more likely to judge the conclusions of agents with female names as being wrong than agents with male names. We used a logistic regression to test the effect of knowledge condition and gender condition on ratings of Right/Wrong. Contrary to hypotheses, there was no main effect (χ^2 = 1.25, p = .26, OR = .77) or moderating effect (p = .99) of gender name. Male and female agents were equally likely to be rated as 'right' in all knowledge conditions. There was only a significant main effect of knowledge condition (χ^2 = 332.89, p < .001). As expected, *post-hoc* tests showed that participants in the knowledge condition were significantly more likely to rate the agent as 'right' (99%) than participants in the Gettier (93%; p = .03, OR = 9.72) and ignorance conditions (5%; p < .001, OR = 2296.67). And, participants in the Gettier condition were significantly more likely to rate the agent as 'right' than participants in the ignorance condition (p < .001, OR = 236.29).

Hypothesis 3: Luck/Ability

We hypothesized that participants would attribute similar amounts of luck (low levels) in the true knowledge and ignorance conditions. However, in the more ambiguous Gettier condition, participants would be more likely to attribute luck to agents with female names than with male names. We used an ANOVA to test the effect of knowledge condition and gender condition on ratings of luck and ability. Higher scores on this scale (0-100) indicated that participants judged an agent to be lucky. There were significant main effects of knowledge condition (F[2,352] = 17.97, p < .001, η^2 = .10, see Figure 2) and gender condition (F[1,352] = 5.54, p = 0.02, η^2 = .01; see Figure 3).

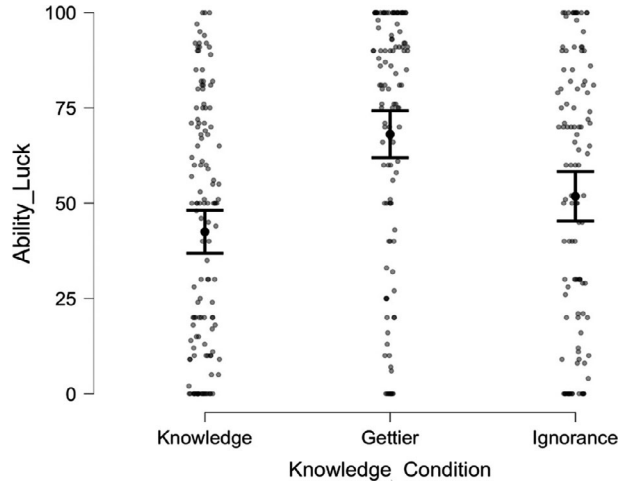

Figure 2. Ability versus Luck rating (sliding scale; 0 – Ability, 100 – luck) as a function of knowledge condition (with standard error bars).

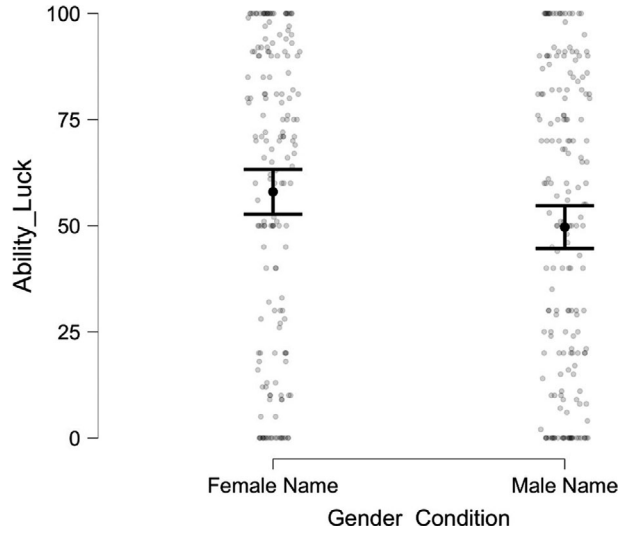

Figure 3. Ability versus Luck rating (sliding scale; 0 – Ability, 100 – luck) as a function of gender condition (with standard error bars).

Participants were significantly more likely to rate the judgments of the agent as due to luck rather than skill in the Gettier condition (M = 68) than in the knowledge (M = 43; $p$tukey < .001) and ignorance conditions (M = 52; $p$tukey < .001). And, participants in the ignorance condition were significantly more likely to rate the judgments of the agent as due to luck rather than skill than in the knowledge condition, $p$tukey = .03. Participants were more likely to rate the knowledge of the agent as due to luck rather than skill in the female name condition (M = 58) than in the male name condition (M = 50).

Secondary analysis: Scenario

One possible explanation for the inconsistencies is that researchers have not used a standardized set of knowledge scenarios; however, researchers have yet to specify what content (or how that content) would lead to different attributions. Thus, we hypothesized that scenario type would not affect knowledge attributions. We used ANOVA and logistic regression to test the effect of scenario condition on the three rating probes: knowledge, right/wrong, and luck/ability. Contrary to hypotheses, results showed that knowledge attributions were not robust to scenario type. Results showed a significant main effect of scenario, ($F[2, 349] = 29.61, p < .001, \eta^2 = .09$) on knowledge attributions. Participants attributed less knowledge to agents in the jewelery scenario ($M = 35$) than agents in the squirrel ($M = 67$; $t = 6.38$, $p$tukey $< .001$) and house scenarios ($M = 61$, $t = 6.81$, $p$tukey $< .001$). There was also a significant scenario condition $\times$ knowledge condition interaction ($F[4, 349] = 2.62$, $p = .04$, $\eta^2 = .02$; see Figure 4) showing that the low knowledge ratings in the jewelery scenario condition were driven largely by the Gettier case judgments ($M = 27$).

Further, there was an effect of scenario condition on ratings of luck/ability. Specifically, results showed a significant knowledge condition $\times$ scenario interaction effect ($F[4, 349] = 4.79$, $p < .001$, $\eta^2 = .05$; see Figure 5). For Gettier cases, the judgement of the agent was more likely to be attributed to luck than ability in the jewelery scenario ($M = 88$) than in the squirrel ($M = 55$, $t = -5.55$, $p$tukey $< .001$) and house ($M = 65$, $t = -6.86$, $p$tukey $< .001$) scenarios. There was no effect of scenario type on judgments of right and wrong.

Exploratory analyses

We conducted two exploratory analyses. First, we examined the effect of participant gender on knowledge attributions. Prior research shows that, in some cases, males are

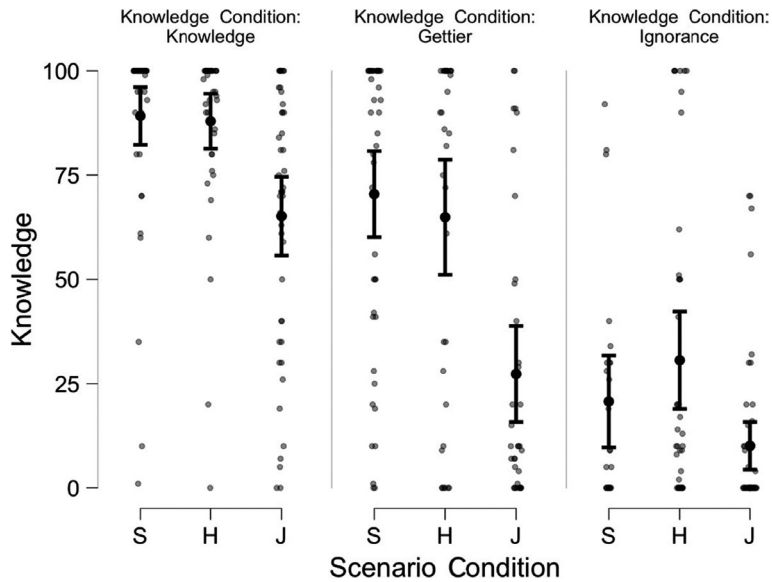

Figure 4. Knowledge rating (0–100) as a function of knowledge condition and scenario condition (with standard error bars). S = Squirrel scenario, H = House Scenario, J = Jewelery Scenario.

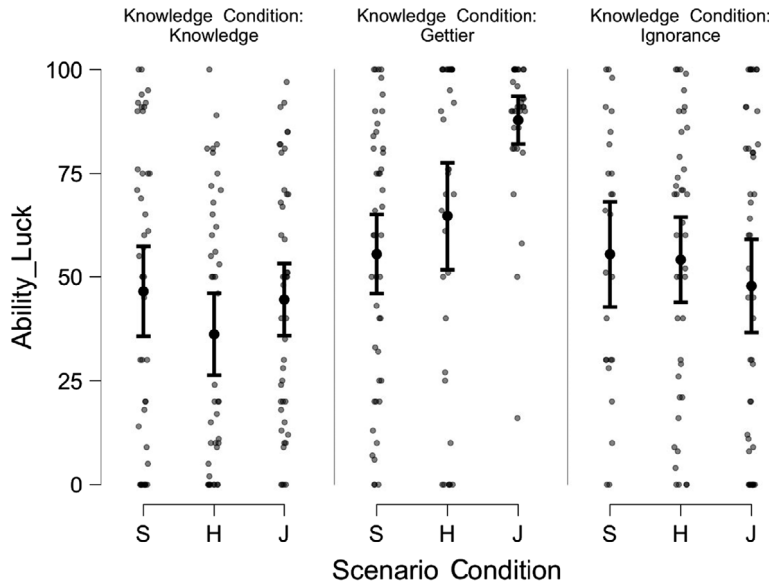

Figure 5. Ability versus Luck rating (sliding scale; 0 – Ability, 100 – luck) as a function of knowledge condition and scenario condition (with standard error bars). S = Squirrel scenario, H = House Scenario, J = Jewelery Scenario.

more likely than females to exhibit a gender bias against females (Dutt et al., 2016; Moss-Racusin et al., 2012; ; Souchon, Livingstone, & Maio, 2013). However, other studies demonstrate that both males and females exhibit a gender bias against women (Basow, 1995; Fan et al., 2019). We conducted an exploratory analysis testing the effect of participant gender on knowledge probe ratings. Results showed that there was not a significant effect of participant's gender on knowledge attributions, judgments of right/wrong, or judgments of luck/ability.

In the second exploratory analysis, we tested the hypotheses in each scenario separately (i.e., squirrel, house, and jewelery, respectively). Given the variability among the scenarios, it is possible that combining them may have masked effects of gender and knowledge condition.

Overall, a similar pattern of results was found when examining each scenario separately. There were two exceptions. First, results showed that in the house and jewelery conditions, respectively, there was not a main effect of knowledge condition on ratings of Right/Wrong. Second, in the squirrel condition, there was not a main effect of knowledge condition or gender condition on ratings of ability/luck (both p's $> .30$; note that the pattern of findings was in the same direction as the primary analysis; the lack of significance was likely due to much lower power). This null finding is consistent with the results of the comprehension check, which suggests that participants had particular difficulty understanding the squirrel scenario (e.g., the squirrel – ignorance condition accounted for the highest number of comprehension errors).

Discussion

The purpose of this study was to examine the effect of gender on knowledge attributions.

We hypothesized that agents with traditional female names would be viewed as less knowledgeable, their answers viewed as less correct, and their beliefs attributed more to luck when compared to agents with traditional male names, particularly in Gettier situations. Contrary to our hypothesis, results showed that male and female agents were judged as having similar amounts of knowledge and as being equally likely to have the correct answer across knowledge conditions. What differed, however, was how participants viewed that knowledge. The knowledge of female agents was more likely to be attributed to luck (vs. ability) than was the knowledge of male agents. This gender bias in luck/ability attributions was found across knowledge conditions, not just for the ambiguous Gettier case.

These findings were not predicted a priori and, perhaps, indicate a societal shift in gender biases (Basow, 1995). Participants attributed equal amounts of knowledge to male and female agents, even for Gettier cases in which there is some latitude to exhibit bias. Similarly, participants were equally likely to rate the judgments of male and female agents as being correct across all conditions. These findings suggest that there may be a reduction in more overt forms of gender bias in which men are judged more positively than women despite equivalent knowledge credentials. This seems to be a positive sign for gender equality. However, we still detected a gender bias. Although male and female agents were judged as equally knowledgeable and correct in their judgements, there was still a bias in how participants viewed that knowledge. The knowledge of female agents, compared to male agents, was more likely to be attributed to luck than ability.

Our results corroborate Deaux and Emswiller's (1974) classic research showing that when men and women perform a traditionally masculine task, women who are successful are viewed as lucky, whereas men who are successful are viewed as skilled (see also Espinoza et al., 2014). Forty-five years later it appears this gender bias in luck/ability attributions still exists. Our study extends this work by showing that the bias to attribute luck to women may not be confined to performance on a task, but also encompasses one's body of knowledge. Further, our results suggest that this bias may not be specific to stereotypically masculine contexts. We found that the knowledge of women agents was more likely to be attributed to luck than the knowledge of male agents regardless of context (ecologist, house viewer, jewelery buyer). This indicates that the gender bias in attributing luck versus ability may be more pervasive than previously thought.

If women's successes and knowledge are not being treated as true knowledge or ability, then this may have a number of real-world implications, most notably in employment and school settings (e.g., Espinoza et al., 2014). For example, our results help explain a recent study showing that when intellectual ability was required in a job description, potential employers selected male candidates at a higher rate than female candidates (Bian, Leslie, & Cimpian, 2018). Ability and luck attributions can also affect classroom dynamics. If teachers attribute female success and knowledge to luck rather than ability, then it may become part of the hidden curriculum, which could affect teacher-student interactions and students' educational trajectories. Indeed, teachers' perceptions of their students' abilities is correlated with students' long-term academic success, suggesting that gender bias in the classroom could be detrimental to students' learning and development (De Boer, Bosker, & van der Werf, 2010). This is particularly true for student success in STEM. Research shows that despite females performing as well if not better in math and science, females are consistently underrepresented in STEM majors (Ganley, George, Cimpian, & Makowski, 2018). Further, universities are less likely to hire a female candidate than a male candidate with identical credentials, and when the

female candidate is chosen, her proposed salary is lower than that of the male candidate (Moss-Racusin et al., 2012).

Research suggests that gender biases in attributions of luck versus ability appear to be internalized over time. For example, Stipek and Gralinski (1991) found that when students were asked about their own successes and failures on a math assessment, male students were more likely to attribute their successes to ability than female students. Additionally, when male and female students were asked to predict their performance, male students predicted higher scores than their female counterparts did (Sieverding & Koch, 2009; Stipek & Gralinski, 1991). It appears the internalization of gender biases in luck versus ability attributions (both external and internal forms) may be one of the reasons that women drop out of scientific fields and are consistently paid lower salaries than men (Jones & Urban, 2013).

It is also important to consider that the gender bias in luck/ability attributions is not just hurting females, but also helping males. If male's knowledge is attributed to ability, then this may lead them to work harder (as they control their success) and continue in STEM fields even when they are not as qualified or knowledgeable as their female counterparts. This 'enriching effect' for males might serve to maintain or even increase gender biases over time. Thus, gender biases in ability and luck have the potential to contribute to a self-fulfilling prophecy.

How can society combat this kind of gender bias in luck and ability attributions? First, it may be important to inform teachers and other educators that this bias exists (Devine et al., 2017). Our results indicate that gender bias is relatively covert in nature. We did not find differences in the overall level of knowledge attributed to male and female agents, but rather a bias in how that knowledge was perceived. Thus, people may not be aware that they exhibit this bias. Research indicates that educating people about potential implicit biases is one of three key factors in reducing such biases. For example, Devine and colleagues (2017) have created an empirically supported intervention for reducing implicit bias in hiring. The intervention consists of three components: recognizing participants' own unintentional biases, understanding the consequences of these biases, and learning strategies to reduce unintentional biases. A recent test of this intervention showed that participants who completed the workshop hired 18% more women compared to hiring rates before the workshop (Devine et al., 2017).

Teachers could deter gender bias in grading by having students show their work and by removing names during grading. For example, math assessments in which students have to show their work or explain their reasoning rather than simply circling an answer may make it harder to attribute a student's knowledge to luck rather than ability. Similarly, when grading assignments, it may be important for teachers to remove students' names from their work in order to ensure impartiality. The removal of names may also help fight gender bias in hiring decisions. For example, the Ecology and Evolutionary Biology Department at the University of Connecticut recently changed their hiring process by eliminating any information that could reveal a candidate's gender (Jones & Urban, 2013). Results showed that they were successful in conducting a gender-blind evaluation with up to 60% of their candidates. This work suggests that gender-blind hiring could be one way to mitigate bias in academia (Jones & Urban, 2013). The practice of removing names and gendered information may prevent a candidate's accomplishments from being attributed to luck rather than skill on the basis of gender.

A secondary purpose of the current study was to determine if the JTB methodology was robust to changes in scenario. Although our findings supported JTB theory and the uniqueness of the Gettier case (participants attributed less knowledge in Gettier cases

than in true knowledge cases), we showed that results were not robust to scenario. Scenario type had a significant effect on comprehension, and thus, which participants were included in the data analysis. The squirrel/ignorance scenario accounted for one third of the incorrect responses to the comprehension probe, indicating that there was difficulty understanding this case. Moreover, the gender of the agent in the scenario affected comprehension. Participants assigned to scenarios with female agents were more likely to fail the comprehension probe than those assigned to scenarios with male agents. Again, this means that prior studies that used scenarios with female agents likely had higher attrition due to failed comprehension than studies that used scenarios with male agents. These results show that scenario and the gender of the agent are leading to selective attrition. Scenario also affected knowledge attributions. Participants were less likely to attribute knowledge and ability to agents in the jewelery scenario, particularly in the Gettier condition. The use of the jewelery scenario needs to be considered when interpreting the results of past studies testing JTB theory.

To our knowledge, our study is the first to show that comprehension and knowledge attributions are not robust to scenario type and the agent's gender. This may help to explain inconsistencies in the literature, especially given small sample sizes used in prior studies, which are more easily skewed by methodological differences such as the ratio of male/female agents or the specific scenario type that was used (Nagel et al., 2013; Turri et al., 2015). These confounds need to be considered when interpreting results from JTB studies. It is possible that the differences found among knowledge conditions may be an artefact of poor comprehension rather than actual differences in knowledge attributions. The Gettier and ignorance scenarios are more difficult to understand than knowledge scenarios (comprehension rates were 10% lower for Gettier and ignorance cases than for true knowledge cases). Participants may be rating agents in these conditions (Gettier and ignorance) as possessing less knowledge simply because they, themselves, are confused about what is true knowledge in those scenarios. It will be important to address this confound by creating a standardized set of scenarios in which the knowledge conditions (knowledge, Gettier, and ignorance) are equated on comprehension scores (as well as agent gender). That said, it may be difficult to manipulate the 3 JTB requirements within each knowledge condition so that they are all equally comprehensible and relatable. For example, it is common for a person to see a new house (true knowledge example), but one rarely, if ever, encounters a fake house (Gettier example). Ultimately, it may be necessary to develop alternative strategies for testing JTB theory that do not rely on written stories (e.g., short videos).

The study had both strengths and limitations. A significant strength of this study was that it was pre-registered with specific hypotheses. The study also had a large sample size ensuring a high level of power (.93) to detect differences among experimental conditions. Additionally, the study included the use of multiple gendered names and multiple scenarios to rule out possible confounds of a specific name or specific scenario driving the results. The primary limitation of the study was the use of a U.S. sample (to ensure common knowledge of gendered names). It remains unclear if the results will generalize to other countries or cultures.

In conclusion, this study contributes to the knowledge base on both gender bias and JTB theory. Results showed that women and men are considered to be equally knowledgeable, but knowledge is more likely to be attributed to luck (than ability) for women than men. This work indicates that while overt gender biases in attributions about knowledge may be fading, there continue to be more covert biases that can still have real-world implications. The results also highlight the need to standardize the scenarios and

gender of the agents in this area of study. It is critical to consider both the scenario and gender of the agent when interpreting the results of studies testing JTB theory as these affect important outcome variables.

Conflicts of interest

The authors declare no conflicts of interest and no funding agencies.

Author contribution

All authors contributed equally to all aspects of manuscript preparation; author order was determined alphabetically. Natalie G. Disher (Conceptualization; Formal analysis; Investigation; Methodology; Project administration; Writing – original draft; Writing – review & editing) Antonio L. Guerra (Conceptualization; Data curation; Formal analysis; Investigation; Methodology; Writing – original draft; Writing – review & editing) Gerald Haeffel (Conceptualization; Data curation; Formal analysis; Investigation; Methodology; Project administration; Supervision; Visualization; Writing – original draft; Writing – review & editing)

Data availability statement

Data can be found here: https://osf.io/b3jfw/

References

Alexander, D., & Andersen, K. (1993). Gender as a factor in the attribution of leadership traits. *Political Research Quarterly*, *46*, 527–545. https://doi.org/10.1177/106591299304600305

Basow, S. A. (1995). Student evaluations of college professors: When gender matters. *Journal of Educational Psychology*, *87*, 656–665. https://doi.org/10.1037/0022-0663.87.4.656

Bian, L., Leslie, S.-J., & Cimpian, A. (2018). Evidence of bias against girls and women in contexts that emphasize intellectual ability. *American Psychologist*, *73*, 1139–1153. https://doi.org/10.1037/amp0000427

Bigelow, L., Lundmark, L., Parks, J. M., & Wuebker, R. (2012). Skirting the Issues. *Journal of Management*, *40*, 1732–1759. https://doi.org/10.1177/0149206312441624

Bosak, J., & Sczesny, S. (2011). Gender bias in leader selection? Evidence from a hiring simulation study. *Sex Roles*, *65*(3–4), 234–242. https://doi.org/10.1007/s11199-011-0012-7

De Boer, H., Bosker, R. J., & van der Werf, M. P. C. (2010). Sustainability of teacher expectation bias effects on long-term student performance. *Journal of Educational Psychology*, *102*(1), 168–179. https://doi.org/10.1037/a0017289

Deaux, K., & Emswiller, T. (1974). Explanations of successful performance on sex-linked tasks: What is skill for the male is luck for the female. *Journal of Personality and Social Psychology*, *29* (1), 80–85. https://doi.org/10.1037/h0035733

Devine, P. G., Forscher, P. S., Cox, W. T. L., Kaatz, A., Sheridan, J., & Carnes, M. (2017). A gender bias habit-breaking intervention led to increased hiring of female faculty in STEMM departments. *Journal of Experimental Social Psychology*, *73*, 211–215. Available from: https://doi.org/10.1016/j.jesp.2017.07.002

Dutt, K., Pfaff, D., Bernstein, A. et al. (2016). Gender differences in recommendation letters for postdoctoral fellowships in geoscience. *Nature Geosci 9*, 805–808. https://doi.org/10.1038/ngeo2819

Espinoza, P., Da Luz, A., Fontes, A. B., & Arms-Chavez, C. J. (2014). Attributional gender bias: Teachers' ability and effort explanations for students' math performance. *Social Psychology of Education*, *17*(1), 105–126. https://doi.org/10.1007/s11218-013-9226-6

Fan, Y., Shepherd, L. J., Slavich, E., Waters, D., Stone, M., Abel, R., & Johnston, E. L. (2019). Gender and cultural bias in student evaluations: Why representation matters. *PLoS One*, *14*, e0209749. https://doi.org/10.1371/journal.pone.0209749

Ganley, C. M., George, C. E., Cimpian, J. R., & Makowski, M. B. (2018). Gender equity in college majors: Looking beyond the STEM/non-STEM dichotomy for answers regarding female participation. *American Educational Research Journal*, *55*, 453–487. https://doi.org/10.3102/0002831217740221

Gettier, E. L. (1963). Is justified true belief knowledge? *Analysis*, *23*, 121–123. https://doi.org/10.1093/analys/23.6.121

Hall, B. F., Wagge, J. R., Pfuhl, G., Stieger, S., Vergauwe, E., Ijzerman, H., ... Stephen, I. (2018). Accelerated CREP – RRR: Turri, Buckwalter, & Blouw (2015). doi: https://doi.org/10.31234/osf.io/zeux9

Heilman, M. E. (2001). Description and prescription: How gender stereotypes prevent women's ascent up the organizational ladder. *Journal of Social Issues*, *57*, 657–674. https://doi.org/10.1111/0022-4537.00234

Jones, C. S., & Urban, M. C. (2013). Promise and pitfalls of a gender-blind faculty search. *BioScience*, *63*, 611–612. https://doi.org/10.1525/bio.2013.63.8.3

Knobloch-Westerwick, S., Glynn, C. J., & Huge, M. (2013). The Matilda effect in science communication. *Science Communication*, *35*, 603–625. https://doi.org/10.1177/1075547012472684

Moss-Racusin, C. A., Dovidio, J. F., Brescoll, V. L., Graham, M. J., & Handelsman, J. (2012). Science faculty's subtle gender biases favor male students. *Proceedings of the National Academy of Sciences*, *109*(41), 16474–16479. https://doi.org/10.1073/pnas.1211286109

Nagel, J., Juan, V. S., & Mar, R. A. (2013). Lay denial of knowledge for justified true beliefs. *Cognition*, *129*, 652–661. https://doi.org/10.1016/j.cognition.2013.02.008

Peer, E., Brandimarte, L., Samat, S., & Acquisti, A. (2017). Beyond the Turk: Alternative platforms for crowdsourcing behavior research. *Journal of Experimental Social Psychology*, *70*, 153–163. https://doi.org/10.1016/j.jesp.2017.01.006

Proudfoot, D., Kay, A. C., & Koval, C. Z. (2015). A gender bias in the attribution of creativity. *Psychological Science*, *26*, 1751–1761. https://doi.org/10.1177/0956797615598739

Sieverding, M., & Koch, S. C. (2009). (Self-)Evaluation of computer competence: How gender matters. *Computers and Education*, *52*, 696–701. https://doi.org/10.1016/j.compedu.2008.11.016

Souchon, N., Livingstone, A. G., & Maio, G. R. (2013). The influence of referees' expertise, gender, motivation, and time constraints on decisional bias against women. *Journal of Sport and Exercise Psychology*, *35*, 585–599. https://doi.org/10.1123/jsep.35.6.585

Steinpreis, R. E., Anders, K. A., & Ritzke, D. (1999). The impact of gender on the review of the curricula vitae of job applicants and tenure candidates: A national empirical study. *Sex Roles*, *41*, 509–528. https://doi.org/10.1023/A:1018839203698

Stipek, D. J., & Gralinski, J. H. (1991). Gender differences in children's achievement-related beliefs and emotional responses to success and failure in mathematics. *Journal of Educational Psychology*, *83*, 361–371. https://doi.org/10.1037/0022-0663.83.3.361

Turri, J. (2016). Knowledge and assertion in "Gettier" cases. *Philosophical Psychology*, *29*, 759–775. https://doi.org/10.1080/09515089.2016.1154140

Turri, J., Buckwalter, W., & Blouw, P. (2015). Knowledge and luck. *Psychonomic Bulletin & Review*, *22*, 378–390. https://doi.org/10.3758/s13423-014-0683-5

Received 30 July 2020; revised version received 5 January 2021

Appendix :

Scenarios

Squirrel scenario

Everyone Reads: (Agent name) is an ecologist collecting data on red-speckled ground squirrels in Canyon Falls National Park. The park is divided into ten zones, and today (agent name) is working in Zone 3. While scanning the river valley with his/her binoculars (agent name), sees a small, bushy-tailed creature with distinctive red markings on its chest and belly. The red-speckled ground squirrel is the only native species with such markings. (Agent name) records in his/her journal, 'At least one red-speckled ground squirrel in Zone 3 today'.

Knowledge

Ecologists are unaware that a complex network of aquifers recently began drying up in parts of the park. These aquifers carry vital nutrients to the trees and other forms of plant life that support the squirrels. And the aquifers in the river valley running through Zone 3 are no exception. The animal (agent name) is looking at is indeed a thirsty red-speckled ground squirrel.

Gettier

Ecologists are unaware that a non-native species of prairie dog recently began invading the park. These prairie dogs also have red markings on their chest and belly. When these prairie dogs tried to invade Zone 3, the red-speckled ground squirrels were unable to completely drive them away. Still, the animal (agent name) is looking at is indeed a red-speckled ground squirrel.

Ignorance

Ecologists are unaware that a non-native species of prairie dog recently began invading the park. These prairie dogs also have red markings on their chest and belly. When these prairie dogs tried to invade Zone 3, the red-speckled ground squirrels were unable to completely drive them away. And, the animal (agent name) is looking at is indeed one of the prairie dogs.

House Scenario

Everyone Reads: (Agent name) is driving through the countryside with his/her young son Andrew. Along the way he/she sees numerous objects and points them out to his/her son. 'That's a cow, Andrew', (agent name) says, 'and that over there is a house where farmers live'. (Agent name) has no doubt about what the objects are.

Knowledge

What (agent name) and Andrew do not realize is the area they are driving through was recently hit by a very serious tornado. This tornado did not harm any of the animals, but did destroy most buildings. In an effort to maintain the rural area's tourist industry, local

townspeople rebuilt new houses in the place of the destroyed houses. These new houses look exactly like the old houses and can be used as actual housing.

Having just entered the tornado-ravaged area, (agent name) notices the many houses lining the roads. When he/she tells Andrew 'That's a house', the object he/she sees and points at is a real house.

Gettier

What (agent name) and Andrew do not realize is the area they are driving through was recently hit by a very serious tornado. This tornado did not harm any of the animals, but did destroy most buildings. In an effort to maintain the rural area's tourist industry, local townspeople built fake houses in the place of destroyed houses. These fake houses look exactly like real houses from the road, but are only for looks and cannot be used as actual housing.

Having just entered the tornado-ravaged area, (agent name) has not yet encountered any fake houses. When he/she tells Andrew 'That's a house', the object he/she sees and points at is a real house that has survived the tornado.

Ignorance

What (agent name) and Andrew do not realize is the area they are driving through was recently hit by a very serious tornado. This tornado did not harm any of the animals, but did destroy most buildings. In an effort to maintain the rural area's tourist industry, local townspeople built fake houses in the place of destroyed houses. These fake houses look exactly like real houses from the road, but are only for looks and cannot be used as actual housing.

When he/she tells Andrew 'That's a house', the object he/she sees and points at is actually a fake house that was built after the tornado.

Jewelery Scenario

Everyone Reads: (Agent name) is shopping for jewelery.

Knowledge

He/she goes into a nice-looking store, and spends some time looking at various different displays. He/she tells the salesperson that he/she is looking for a simple diamond watch with a classic design. He/she always likes to try things on before he/she makes up his/her mind about them, and he/she asks the salesperson to show him/her quite a wide variety of different items, which he/she brings out for him/her one tray at a time. (Agent name) selects a diamond watch from a tray marked 'Diamond Watches'. 'What a lovely diamond!' he/she says as he/she tries it on.

Gettier

He/she goes into a nice-looking store, and selects a diamond watch from a tray marked 'Diamond Watches'. 'What a lovely diamond!' he/she says as he/she tries it on. (Agent name) could not tell the difference between a real diamond and a cubic zirconium fake just by looking or touching. In fact, this particular store has a very dishonest employee who has

been stealing real diamonds and replacing them with fakes; in the tray (agent name) chose, almost all of the watches had cubic zirconium stones rather than diamonds (but the one he/she chose happened to be real).

Ignorance

He/she goes into a nice-looking store. He/she looks at several displays, then selects a watch from a tray marked 'Diamond Watches'. 'What a lovely diamond!' he/she says as he/she tries it on. (Agent name) could not tell the difference between a real diamond and a cubic zirconium fake just by looking or touching. In fact, this particular store has a dishonest employee who has been stealing real diamonds and replacing them with fakes; in the tray (agent name) chose from, all of the watches – including the one he/she tried on – had cubic zirconium stones rather than diamonds.

Probes

Squirrel probes

1. (Agent name)_____that there is at least one red-speckled ground squirrel in Zone 3 today; Only Believes <------> Knows
2. (Agent name) is looking at a_____; *ground squirrel/prairie dog*.
3. It is_____for (agent name) to think that he/she is looking at a red-speckled ground squirrel; Unreasonable <-------> Reasonable.
4. (Agent name) got the_____answer because of his/her_____; *right/wrong*; (In)Ability <-------> Good/Bad Luck

House probes

1. (Agent name)_____that he/she is pointing at a real house; Only Believes <------> Knows
2. (Agent name) is pointing at a_____house; *real/fake*
3. It is_____for (agent name) to think that he/she is pointing at a real house; Unreasonable <------> Reasonable
4. (Agent name) got the_____answer because of his/her_____; *right/wrong*; (In) Ability <---------> Good/Bad Luck

Jewelery probes

1. (Agent name)_____that he/she chose a watch made of diamonds; Only Believes <------> Knows
2. (Agent name) chose a watch made of_____; *cubic zirconium stones/diamonds*
3. It is_____for (agent name) to think that he/she chose a watch made of diamonds; Unreasonable <-------> Reasonable
4. (Agent name) got the_____answer because of his/her_____; *right/wrong*; (In)Ability <---------> Good/Bad Luck