Title
Considering Alternative Outcomes of Research: Does Knowing the Actual Outcome Create Bias?

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 45(45)

Authors
Munnich, Edward
Adolfo, Marci
Divinagracia, Erica
et al.

Publication Date
2023

Peer reviewed

Considering Alternative Outcomes of Research: Does Knowing the Actual Outcome Create Bias?

Edward Munnich (emunnich@usfca.edu)
Marci Adolfo (mbadolfo2@dons.usfca.edu)
Erica Divinagracia (egdivinagracia@dons.usfca.edu)
Sabrina Ibrahim (sdibrahim@dons.usfca.edu)
Sydney Laxson (srlaxson@dons.usfca.edu)
Joshua Condry (jjcondry@dons.usfca.edu)
Psychology Department, University of San Francisco, 2130 Fulton St.,
San Francisco, CA, 94117 United States

Abstract

Learning a research outcome in class or the media may bias people towards that outcome (hindsight bias), and receiving an explanation may accentuate bias (explanation bias), both of which could hinder understanding of the necessity of replication. We tested whether providing outcomes and explanations of research findings increased difficulty of explaining alternative outcomes, and, if so, whether people were less surprised by the presented findings, and found them more likely to replicate. Amazon Mechanical Turk (AMT) workers and introductory psychology students were randomly assigned to do one of the following: 1. Read details of four psychological studies without their outcomes, 2. additionally receive the outcomes, 3. additionally receive explanations of outcomes. We did not find reliable effects on difficulty of explaining alternative outcomes, and found little evidence for hindsight or explanation biases. We speculate that explaining alternative outcomes immediately after considering the actual outcomes may have debiased our participants.

Keywords: hindsight bias; explanation bias; surprise; replication; teaching psychology; science journalism

In classes and media coverage about research, people often read "Researchers found X, due to Y" offering both the outcome (X) and an explanation (Y) of the research. However, a body of work in cognitive science suggests that presenting findings in this way might induce biases that make it difficult to imagine different outcomes or think of different explanations than those one learned about. Given the number of psychological findings taught in classes and reported in the media that have failed to replicate (Open Science Collaboration, 2015), it is crucial to remain appropriately skeptical of the results of single research studies. In this paper, we consider whether learning an outcome and having an explanation for it makes it more difficult to consider possible alternative outcomes, and if so, whether this leads to biases towards the presented outcome.

Slovic and Fischhoff (1977) found hindsight bias in consideration of experimental outcomes, among participants who received the actual outcomes (and this finding was replicated by Chen et al., 2021). Specifically, after all participants read about the designs of experiments, those who were shown the results of those experiments (Hindsight group) found the results less surprising, and predicted them to be more likely to replicate, than did participants who were asked about the results hypothetically (Foresight group). Moreover, even when Slovic and Fischhoff asked Hindsight participants to explain how the experiment could have had the *opposite* results, hindsight bias persisted, albeit to a lesser degree.

Beyond hindsight bias, numerous studies have demonstrated explanation bias–when participants have an explanation for a phenomenon in mind, it is difficult for them to conceive of different outcomes. For example, Anderson et al. (1980) found that participants persisted in believing either that risk-taking benefited or harmed firefighters after explaining data consistent with the respective conclusions, even after being told that the data were fake. This suggested that the act of explaining—even when that which one explained is known to be unreliable—has an enduring effect on one's beliefs. Explanation bias may be even more powerful when one is provided with an explanation. In this vein, Wong (1995) considered whether explanations of outcomes affected participants' ratings of how obvious those outcomes seemed. Participants read about 12 different studies on teaching methods; for each study, they received either the actual outcome or an opposite outcome, and either received an explanation of that outcome or no explanation. She found that providing explanations increased ratings of obviousness of those outcomes. One possible mechanism for this explanation is suggested by Maguire et al. (2011), who found that in considering a surprising outcome (e.g., how a person could have overslept on the day of an important meeting), participants who received a plausible explanation (e.g., a power outage prevented their alarm from going off) found the outcome less surprising than those who came up with their own explanation. Although Maguire et al. did not test participants' ability to consider an alternative outcome, they argued that the certainty of being told why an outcome occurred was the source of reduced surprise. As such, explanation bias might increase when explanations are provided rather than generated by participants.

Building on findings that explaining possible alternative outcomes can reduce bias, other research demonstrates that alternative outcomes must be relatively easy to imagine for people to take them into account, and suggests a mechanism by which debiasing occurs. Sanna et al. (2002) hypothesized that being asked to provide a small number of reasons for an outcome allows participants to report reasons that easily

In M. Goldwater, F. K. Anggoro, B. K. Hayes, & D. C. Ong (Eds.), *Proceedings of the 45th Annual Conference of the Cognitive Science Society.*

come to mind, but being asked for many reasons for the outcome requires them to mentally search for more obscure reasons, and the difficulty of this search leads them to consider the outcome less plausible. As hypothesized, Sanna et al. found that participants who were asked for two thoughts about an alternative outcome in a war between the Gurkha and British participants considered the alternative more likely, but those who were asked for 10 thoughts about the alternative outcome considered the alternative *less* likely. Similarly, Hirt et al. (2004) found that debiasing was moderated by task difficulty and need for cognition–participants who were high in need for cognition provided judgments that were less biased towards focal outcomes when they were asked to consider alternative outcomes for a NBA basketball division that were easier to mentally simulate (e.g., a different strong team, as opposed to a different weak team, winning a division). Moreover, this debiasing effect transferred to judgments in other domains (NFL football or the best sitcoms), suggesting a mechanism for debiasing: For those high in need for cognition, considering an easy-to-imagine alternative in one domain induced a mental simulation mindset that debiased judgments across domains. Taken together, these findings suggest that manipulations that increase the difficulty of imagining an alternative outcome would reduce how likely one thinks that outcome is.

Our current project combines elements of both Slovic and Fischhoff's (1977) and Wong's (1995) studies to test for effects of providing outcomes of research studies and explanations of those outcomes. Our pre-registered hypothesis (https://osf.io/v8j5s) was that providing focal outcomes would make it more difficult for participants to imagine alternative outcomes for research studies, and that difficulty of imagining alternatives would be further heightened by additionally providing plausible explanations for the focal outcomes. If so, we predicted the following:

1. Hindsight bias: When participants are told the actual outcome of the study (Hindsight condition), they will find it more difficult to imagine an alternative outcome, and will therefore report less surprise and higher likelihood of replication of the actual outcome, compared to those who did not know the actual outcome (Foresight condition).

2. Explanation bias: When participants receive an explanation for the actual outcome (Conjunction condition), they will find it more difficult to imagine an alternative outcome, and will therefore report less surprise and higher likelihood of replication of the actual outcome, compared to those who did not receive an explanation (Hindsight and Foresight conditions).

Experiment 1

Method

Participants and Design 112 Amazon Mechanical Turk workers were paid $1.50 for their time, and were randomly assigned to Foresight, Hindsight, or Conjunction groups.

Materials and Procedure For each of four studies which had been reported in journal articles (Table 1), participants read a brief vignette describing the study in a journalistic format–a headline and brief description of the study. After reading one vignette, participants carried out one of the following tasks for that vignette:

- Foresight participants provided an explanation of why the actual outcome of the study could have been the result, not knowing it was the actual outcome.
- Hindsight participants read the actual outcome of the study and provided an explanation of why it was the result.
- Conjunction participants read *both* the actual outcome of the study and a plausible explanation of that outcome.

Next, all participants did the following: 1. explained an *alternative* possible outcome of the study (e.g., "Can you think of a reason why increased exposure to TV at a young age would lead to a lower risk of language delay?"), 2. rated how difficult it was to explain the alternative outcome on a 5-point Likert scale ranging from 1 (*not at all difficult*) to 5 (*extremely difficult*), 3. rated how surprised they *would be* (Foresight group) or *are* (Hindsight and Conjunction groups) by that outcome of the study, on a 5-point Likert scale ranging from 1 (*not at all surprised*) to 5 (*extremely surprised*), and 3. rated the likelihood that the outcome would be replicated if the same study were conducted again, on a 5-point Likert scale ranging from 1 (*extremely UNLIKELY that the results would be the same*) to 5 (*extremely LIKELY that the results would be the same*). After answering these questions for one vignette, participants proceeded to the next vignette and repeated these steps.

Halfway through the experiment, participants came to an attention check which was identical in format to the actual questions, but instructed them *not* to answer any questions about it. Participants who *did* answer the questions were excluded from further analyses. Our full questionnaire was pre-registered and can be found at https://osf.io/qcfxt, with one change: We substituted the Likert scales described above for sliders.

Table 1: Vignettes Used in Experiments 1 and 2.

Study	Question	Citation
Changing Test Answers	Is it better to change one's answers on a test or stick with one's original answers?	(Bauer et al., 2007)
TV-Language Delay	Does greater TV exposure increase or decrease the rate of children's linguistic development?	(Byeon & Hong, 2015)
Font Style-Learning	Do people learn more with materials in a difficult or an easy-to-read font?	(Diemand -Yauman et al., 2011)
Age-Memory	Does false memory of a crime increase or decrease with aging?	(Aizpurua et al., 2009)

Results

Difficulty of Explaining Alternative Outcome As a manipulation check, we carried out ANOVAs for each vignette on difficulty of explaining the alternative outcome. We found reliable differences in difficulty for TV-Language ($F(2, 70.4)=3.81$, $p= .027$) and for Font Style-Learning ($F(2, 67.6)=3.78$, $p=.028$), but no reliable differences for Changing Answers or Age-Memory (Fs<2.41, ps>.05). However the effects were not in the expected direction: We had anticipated that providing the actual outcomes (Hindsight) and additionally providing explanations (Conjunction) would lead participants to mentally simulate the actual outcome, and have more difficulty in imagining an alternative outcome; instead Foresight participants' mean difficulty ratings were higher than those of Hindsight participants for all but one vignette (Age-Memory, which was a non-significant difference), and higher than those of Conjunction participants for all items (Figure 1).

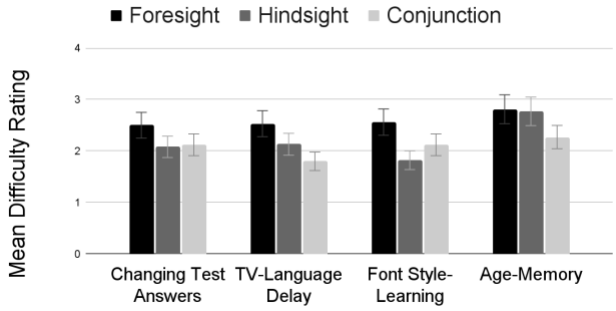

Figure 1: Experiment 1 Mean Difficulty of Explaining Alternative Outcomes.

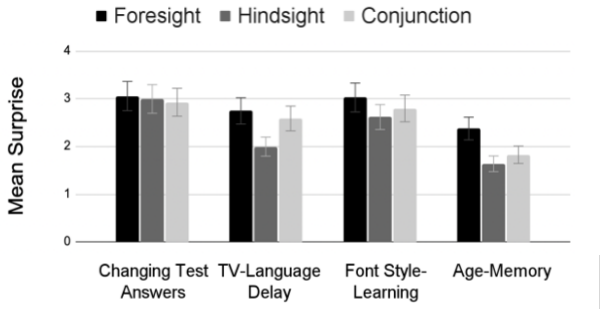

Figure 2: Experiment 1 Surprise Ratings for Actual Outcomes.

Surprise and Likelihood of Replication Next, we carried out ANOVAs to test any effects of our manipulation on surprise and likelihood of replication of the actual outcome. However, since our manipulation check failed to show the expected effects on difficulty of explaining an alternative outcome across our groups—if anything, results trended in the opposite of the expected direction—we no longer expected to find hindsight or explanation biases. We found a significant difference in surprise for TV-Language ($F(2, 70.8)=4.09$, $p=.021$), and a marginally significant difference for Age-Memory ($F(2, 67.0)=3.05$, $p=.054$), but differences in surprise did not reach significance for the other two items (Fs<1.15, ps>.323). As predicted, Foresight participants reported the greatest mean surprise for the actual outcome across all four vignettes (Figure 2). Although several of the effects did not reach significance, this suggests that seeing the actual outcome might reduce surprise for participants. However, given that difficulty trended in the opposite of our expected direction across groups, effects of surprise do not appear to be related to the difficulty of generating an alternative outcome in the way we anticipated.

This trend in surprise did not extend to predictions of likelihood of replication, where we found no significant differences for any of the items (Fs<1.80, ps>.173; Figure 3). Despite the lack of the predicted effects of our manipulation on likelihood of replication, we did find reliably negative Pearson correlations between surprise and likelihood of replication across all four items (Changing Test Answers: $r(110)=-0.37$, $p<.01$; TV-Language: $r(110)=-0.47$, $p<.01$; Font Style- Learning: $r(110)=-0.24$, $p=.01$; Age-Memory: $r(110)=-0.27$, $p=.01$). This supports our predicted relationship between surprise and likelihood of replication: Individuals who were more surprised by the actual outcomes predicted lower likelihood of replication of the studies.

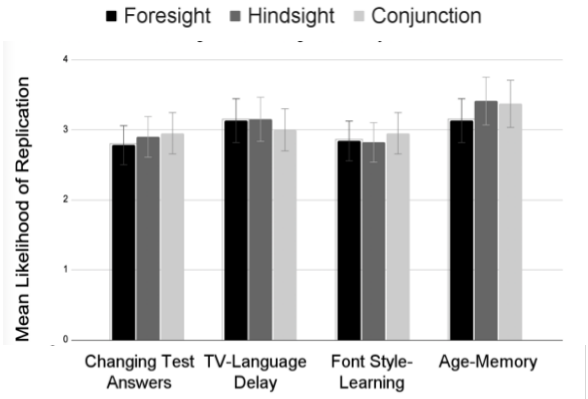

Figure 3: Experiment 1 Likelihood of Replication Ratings for Actual Outcomes.

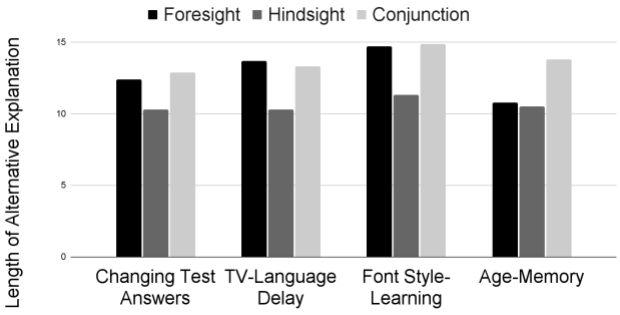

Figure 4: Experiment 1 Length of Alternative Outcome Explanation.

Qualitative Analyses It is possible that Foresight participants provided similar or higher difficulty ratings not because it was actually more difficult for them to think of an alternative outcome, but rather because they were more engaged in the task. If Hindsight and Conjunction participants were convinced that the research could not have turned out a different way, they might have put less effort into their explanations of alternative outcomes, and therefore rated it less difficult than they would if they fully considered an alternative. To test this, we analyzed the length and kinds of responses given by each group. First, Foresight and Conjunction participants provided similar length responses, and Hindsight participants provided the shortest responses across conditions (Figure 4). This is consistent with Hindsight participants putting less effort into their responses, leaving open the possibility that their similar difficulty ratings to Foresight participants could reflect less effort on the task offsetting the difficulty of overcoming a known answer. However, the lack of difference in the length of responses of Foresight and Conjunction participants leaves us no reason to believe that the Conjunction group's difficulty ratings were lower because they put less effort into the task. Digging deeper, we examined the kinds of answers provided by each group and found broad similarity in alternative outcomes across the three groups. For example, for Changing Answers, the most common responses across groups focused on possibilities of overthinking, second-guessing, and gut instinct. For Age-Memory, the most common responses across groups were greater attention, experience, and brain development as major factors for why older people remembered more of a crime scene. Taken together, these results suggest similarity rather than difference across groups in how they thought about alternative outcomes, with the exception that the Hindsight group possibly put less effort into thinking of alternatives.

Discussion

Contrary to our predictions, the manipulation failed to produce differences in difficulty of generating explanations of alternative outcomes. If anything, difficulty of explaining an alternative outcome trended in the direction of greater difficulty for those who did not know the actual outcomes (Foresight group) than for those who were shown the results of studies (Hindsight group) or both results and explanations (Conjunction group). Moreover, we found no consistent evidence that difficulty ratings reflected differences in engagement across the groups. We did find some evidence that actual outcomes were more surprising for the Foresight group than for Hindsight and Conjunction groups, but no effects on predicted likelihood of replication. Nevertheless, we confirmed that there is a relationship between surprise and likelihood of replication on an individual level (despite lack of effects between groups): We found reliable negative correlations between surprise and likelihood of replication across conditions, indicating that when participants were more surprised by an actual outcome, they believed that it was less likely to replicate.

Although we did not find differences in the expected direction for difficulty of explaining alternative outcomes, and only limited support for one of our predictions, it could be that AMT participants find it less difficult to imagine alternative outcomes than undergraduate psychology students, who are presumably more knowledgeable about psychological research and more invested in knowing the correct answers to psychological questions. To test this possibility, and whether a population more engaged in the field would show hindsight bias and explanation bias, we replicated Experiment 1 with students who were taking an introductory psychology class.

Experiment 2

Method

Participants 166 undergraduate introductory psychology students participated for course credit.

Design, Materials, and Procedure were the same as in Experiment 1.

Results

Difficulty of Explaining Alternative Outcome As for Experiment 1, we conducted ANOVAs to investigate differences in difficulty of explaining alternative outcome, and did not find any significant differences on difficulty of generating an alternative outcome (Fs<2.55, ps>0.05). However, unlike Experiment 1, participants' mean ratings of difficulty of generating alternative outcomes trended in the predicted direction: Foresight participants' ratings were numerically lower than Conjunction participants across all four vignettes, and lower than Hindsight participants in all but Age-Memory (Figure 5).

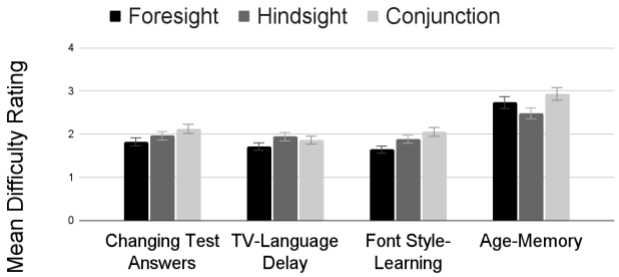

Figure 5: Experiment 2 Mean Difficulty of Explaining Alternative Outcomes.

Surprise and Likelihood of Replication Next, we conducted ANOVAs to test whether our manipulation affected surprise and likelihood of replication judgements. Based on the lack of significant differences in difficulty of explaining alternative outcomes, we would not expect to find significant differences in surprise or likelihood of replication. Accordingly, we did not find reliable effects on surprise for any of the four items (Fs<2.75, ps>0.068). Unexpectedly,

Conjunction participants' mean surprise ratings were numerically *greater* than those of Foresight participants across all four items (Figure 6): Those who received explanations for the actual outcome were more surprised by it than those who received neither the outcomes nor explanations: Changing Test Answers (Foresight: *M*=2.68; Conjunction: *M*=3.03), TV-Language (Foresight: *M*=2.39; Conjunction: *M*=2.40), Font-Style Learning (Foresight: *M*=2.63; Conjunction: *M*=3.00); Age-Memory (Foresight: *M*=1.48; Conjunction: M=1.67).

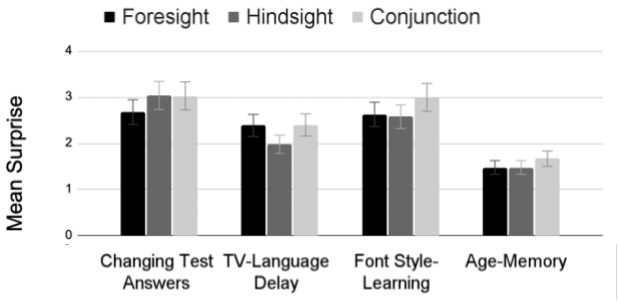

Figure 6: Experiment 2 Surprise Ratings for Actual Outcomes.

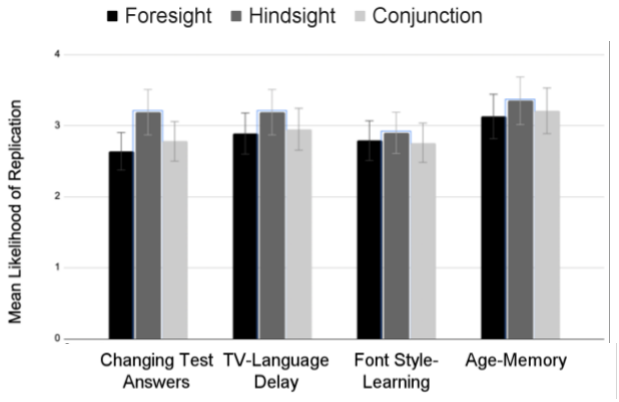

Figure 7: Experiment 2 Likelihood of Replication Ratings for Actual Outcomes.

Analyses of likelihood of replication revealed one significant effect, for Changing Test Answers ($F(2, 105)=8.61$, $p<.001$; here, as predicted, Foresight participants were least likely to believe the outcome would replicate (*M*=2.64), but the order of Conjunction (*M*=2.78), and Hindsight groups (*M*=3.19) were not in the predicted direction. However, we found no other reliable differences for the other three vignettes (*F*s<2.58, *p*s>.081). Nevertheless, as in Experiment 1 and consistent with the predicted relationship between surprise and likelihood of replication on an individual level, those who were more surprised by actual outcomes believed them to be less likely to replicate across three vignettes (TV-Language: $r(164)=-.32$, $p<.001$; Font-Style Learning: $r(164)=-.17$, $p=.03$; Age-Memory: $r(164)=-.52$, $p<.001$), although we found no such relationship for Changing Test Answers ($r(164)=.05$, $p=.56$; Figure 7).

General Discussion

We hypothesized that people experience more difficulty with considering alternative outcomes when they learn the results of experiments (Hindsight group) and even more difficulty when they additionally receive explanations for those outcomes (Conjunction group), and that these would lead to hindsight and explanation bias in terms of surprise and likelihood of replication of the studies. However, in experiments with both AMT and introductory psychology participants, we found little evidence that presenting actual outcomes with or without explanations reliably increased the difficulty of considering alternative outcomes; without systematic differences across tasks in the difficulty of considering alternative outcomes, we would not expect effects of our manipulation on surprise and likelihood of replication, and indeed found very little evidence of this.

In Experiment 1, we tested an AMT population, and contrary to our hypothesis, found that less information about actual outcomes actually led to greater difficulty in explaining alternative outcomes. Although we found some evidence for greater surprise about actual outcomes among those who did not receive those outcomes, we did not find differences in predicted likelihood of replication. Experiment 2 repeated the methods in Experiment 1, but with an introductory psychology population that was presumably more knowledgeable and more invested in "correct" answers to psychological questions. Here, we found trends, albeit in the opposite direction of our predictions and not reaching significance: Foresight participants reported numerically *less* difficulty for explaining an alternative outcome than those who received the actual outcome (Hindsight participants), and those who additionally received an explanation of the actual outcome (Conjunction participants). However, we did not find reliable differences in surprise, and the one significant difference we saw in likelihood of replication was driven by the Hindsight group, again suggesting weak, if any, hindsight bias and no explanation bias.

One possible factor that reduced difficulty of considering alternative outcomes is that participants provided explanations of alternative outcomes immediately after they provided or read explanations for actual outcomes, which might have left participants in all conditions open to alternatives. Regarding our first prediction that there would be hindsight bias, Slovic and Fischhoff's (1977) results indicated that it was reduced in magnitude when participants considered two opposite outcomes of studies (i.e., comparing their foresight participants in their Expt. 1 to hindsight participants in their Expt. 2), compared to when participants only considered one outcome (i.e., comparing their foresight participants in their Expt. 2 to hindsight participants in their Expt. 1). Notably, Chen et al. (2021) only replicated Exp 1 of Slovic and Fischhoff, so their results do not speak to whether bias persists when Hindsight participants consider alternative outcomes. Considering the present study together with Slovic

and Fischhoff's, it could be that hindsight bias is very weak at best, and perhaps can be overcome by simply explaining an alternative possible outcome immediately after one explains the focal outcome. Regarding our second prediction that there would be explanation bias, although Wong (1995) found that participants who were given explanations for outcomes generally rated them more obvious, she did not ask participants to explain how an alternative outcome could have come about. As a result, we do not know if the explanation bias she reported would have withstood explanation of alternative outcomes immediately after receiving an explanation of focal outcomes. As with hindsight bias, an enticing possibility is that explanation bias may be muted or overcome simply by having participants explain an alternative outcome immediately after seeing an explanation of the actual outcome.

Building on the logic of Experiment 2, it could be that populations with greater knowledge of, and investment in psychology–for example, more advanced psychology students and researchers–would show hindsight and explanation biases in experiments like ours. Our present results only speak to concerns about these biases for those with relatively little knowledge of psychology, but those more advanced in the field might find it difficult to explain outcomes that run contrary to other findings they have come to accept as fact. Given that we found negative correlations between surprise and likelihood of replication across both experiments, it may be that when one is surprised by a research outcome that contradicts findings one has come to believe in strongly, one would tend to think the new finding is less likely to replicate. However, even if this is the case, those more advanced in the field are aware of many studies, and may be well justified in thinking of surprising new findings as outliers.

With the caveat that failure to find an effect in two studies does not rule out the possibility of an effect, we speculate that difficulty of imagining an alternative outcome, and therefore hindsight and explanation biases, might be mitigated by as little as asking participants to explain how an alternative outcome could have come about immediately after sharing information about the first outcome. To test this, our next step will be to ask participants in our three conditions about just one outcome for each of our four vignettes; if biases emerge when participants consider only one outcome, taken together with the present results, it would suggest that asking about the alternative outcome immediately after the actual outcome is a simple way to overcome bias. If so, journalists writing for the general public (like our AMT sample), and those teaching undergraduate psychology, could be advised to ask readers/students "how could the opposite have occurred?" when they discuss research outcomes and provide explanations of them.

Acknowledgments

This work was supported by funding from the University of San Francisco Faculty Development Fund. We are grateful for helpful comments from Alyssa Guevara, Sidney Tran, Rose Hoch, and Henna Kalumbu.

References

Aizpurua, A., Garcia-Bajos, E., & Migueles, M. (2009). False memories for a robbery in young and older adults. Applied Cognitive Psychology, 23(2). https://doi.org/10.1002/acp.1461

Anderson, C. A., Lepper, M. R., & Ross, L. (1980). Perseverance of social theories: The role of explanation in the persistence of discredited information. *Journal of Personality and Social Psychology*, *39*(6). https://doi.org/10.1037/h0077720

Bauer, D., Kopp, V. & Fischer, M.R. (2007) Answer changing in multiple choice assessment change that answer when in doubt – and spread the word!. BMC Med Educ 7, 28. https://doi.org/10.1186/1472-6920-7-28

Byeon, H. & Hong, S. (2015). Relationship between Television Viewing and Language Delay in Toddlers: Evidence from a Korea National Cross-Sectional Survey. PLoS ONE 10(3): e0120663. https://doi.org/10.1371/journal.pone.0120663

Chen, J., Kwan, L. C., Ma, L. Y., Choi, H. Y., Lo, Y. C., Au, S. Y., Tsang, C. H., Cheng, B. L., & Feldman, G. (2021). Retrospective and prospective hindsight bias: Replications and extensions of Fischhoff (1975) and Slovic and Fischhoff (1977). *Journal of Experimental Social Psychology*, *96*, Article 104154. https://doi.org/10.1016/j.jesp.2021.104154

Diemand-Yauman, C., Oppenheimer, D. M., & Vaughan, E. B. (2011). Fortune favors the bold (and the italicized): Effects of disfluency on educational outcomes. Cognition, 118(1). https://doi.org/10.1016/j.cognition.2010.09.012

Hirt, E. R., Kardes, F. R., & Markman, K. D. (2004). Activating a mental simulation mind-set through generation of alternatives: Implications for debiasing in related and unrelated domains. *Journal of Experimental Social Psychology*, *40*(3). https://doi.org/10.1016/j.jesp.2003.07.009

Lord, C. G., Lepper, M. R., & Preston, E. (1984). Considering the opposite: A corrective strategy for social judgment. *Journal of Personality and Social Psychology*, *47*(6). https://doi.org/10.1037/0022-3514.47.6.1231

Maguire, R., Maguire, P., & Keane, M. T. (2011). Making sense of surprise: An investigation of the factors influencing surprise judgments. Journal of Experimental Psychology: Learning, Memory, and Cognition, 37(1). https://doi.org/10.1037/a0021609

Open Science Collaboration (2015). Estimating the reproducibility of psychological science. Science, 349(6251).

Sanna, L.J., Schwarz, N. & Small, E.M. (2002). Accessibility experiences and the hindsight bias: I knew it all along versus it could never have happened. *Mem Cogn* 30. https://doi.org/10.3758/BF03213410

Slovic, P., & Fischhoff, B. (1977). On the psychology of experimental surprises. Journal of Experimental

Psychology: Human Perception and Performance, 3(4). https://doi.org/10.1037/0096-1523.3.4.544

Wong, L. Y.-S. (1995). Research on teaching: Process-product research findings and the feelings of obviousness. *Journal of Educational Psychology*, *87*(3), 504–511. https://doi.org/10.1037/0022-0663.87.3.504