Fluency Misattribution and Auditory Hindsight Bias

Daniel M. Bernstein  Ragav Kumar and Michael E. J. Masson  Daniel J. Levitin
Kwantlen Polytechnic University  University of Victoria  McGill University

We conducted three experiments to test the fluency-misattribution account of auditory hindsight bias. According to this account, prior exposure to a clearly presented auditory stimulus produces fluent (improved) processing of a distorted version of that stimulus, which results in participants mistakenly rating that item as easy to identify. In all experiments, participants in an exposure phase heard clearly spoken words 0, 1, 3, or 6 times. In the test phase, we examined auditory hindsight bias by manipulating whether participants heard a clear version of a target word just prior to hearing the distorted version of that word. Participants then estimated the ability of naïve peers to identify the distorted word. Auditory hindsight bias and the number of priming presentations during the exposure phase interacted under-additively in their prediction of participants’ estimates: When no clear version of the target word appeared prior to the distorted version of that word in the test phase, participants identified target words more often the more frequently they heard the clear word in the exposure phase. Conversely, hearing a clear version of the target word at test produced similar estimates, regardless of the number of times participants heard clear versions of those words during the exposure phase. As per Roberts and Sternberg’s (1993) additive factors logic, this finding suggests that both auditory hindsight bias and repetition priming contribute to a common process, which we propose involves a misattribution of processing fluency. We conclude that misattribution of fluency accounts for auditory hindsight bias.

On March 31, 2009, six scientists and a former government official met in L’Aquila, Italy, to ponder the possibility of an earthquake in the region. Although several recent tremors had occurred nearby, the group concluded that it was impossible to predict a major earthquake. Six days later, a 6.3 magnitude quake rocked L’Aquila, killing 300 people. On October 22, 2012, an Italian court sentenced the group to six years in jail for manslaughter. On March 31, 2009, a 6.3 magnitude quake rocked L’Aquila, killing 300 people. On October 22, 2012, an Italian court sentenced the group to six years in jail for manslaughter (Bayen, Pohl, Erdfelder, & Auer, 2007; Bernstein, Erdfelder, Meltzoff, Peria, & Loftus, 2011; Pohl, Bayen, Arnold, Auer, & Martin, 2018; Pohl, Bender, & Lachman, 2002). Traditionally observed using trivia questions, such as “How many feet tall is the Statue of Liberty from base to torch?” and event outcomes, such as elections, sports, and trials, hindsight bias has been observed using visual, auditory, and gustatory judgments (Harley, Carlser, & Loftus, 2004; Lange, Thomas, Dana, & Dawes, 2011; Pohl, Schwarz, Sczesny, & Stahlgberg, 2003; see Bernstein, Abfalg, Kumar, & Ackerman, 2016; Pohl & Erdfelder, 2017 for reviews).

In the auditory hindsight task, participants estimate how many out of 100 of their peers would be able to identify a single distorted word or sentence played once. In a Clear-distorted condition, participants learn the identity of the word (e.g., “barn”) immediately before estimating how many out of 100 of their naïve peers would be able to identify the distorted word (e.g., “bahn”). In a Distorted-only estimation condition, they do not learn the identity of the word; participants hear only the distorted word (e.g., “bahn”) and estimate how many of their naïve peers would be able to identify this word. This task yields robust hindsight bias in that participants reliably estimate that more of their peers would be able to identify the distorted word in the Clear-distorted condition than in the Distorted-only estimation condition (Bernstein, Wilson, Permat, & Meilleur, 2012). As with other types of hindsight bias, auditory hindsight bias occurs when knowing the

We thank Alex Wilson, Jamie Rich, and Bill Peria for help with stimulus preparation, and Bert Sager for helpful suggestions. This work was supported by grants to Daniel M. Bernstein from the Social Sciences and Humanities Research Council of Canada (435-2015-0721) and the Canada Research Chairs Program (950-228407), and by discovery grants to Michael Masson (7910) and Daniel J. Levitin (228175-10) from the Natural Sciences and Engineering Research Council of Canada.

Correspondence concerning this article should be sent to Daniel Bernstein, Department of Psychology, Kwantlen Polytechnic University, 12666-72nd Ave., Surrey, British Columbia V3W 2M8, Canada. E-mail: dbernste@kpu.ca
HINDSIGHT BIAS AND REPETITION PRIMING  

identity of a distorted stimulus makes that stimulus seem more identifiable to a naïve peer than it actually is.

Fluency misattribution in one possible mechanism responsible for hindsight bias (Harley et al., 2004). Fluency is the speed or ease of processing a given stimulus; participants who experience greater fluency for a given item (e.g., because of prior exposure to it) may mistakenly judge that this ease of processing is caused by the item being easy to identify, rather than being advantaged by its previous presentation (see Higham, Neil, & Bernstein, 2017; Werth & Strack, 2003). The fluency-misattribution account of hindsight bias proposes that prior exposure to a stimulus results in people processing that stimulus more fluently, but attributing that fluency to an incorrect source. In a visual hindsight bias task, exposure to a clear face in a Clear-distorted condition results in greater fluency for it when repeated minutes later; assuming the participant considers this fluency to be a product of immediate circumstances (e.g., the obscured face happens to be easy to identify), this would result in underestimating the difficulty of identifying the face (Harley et al., 2004). Calling attention to the clear image—by modifying the task to make the influence of the clear image on judgments obvious—neutralizes the hindsight-bias effect, consistent with other work (Bernstein & Harley, 2007; see Bruner & Potter, 1964; Thomas & Jacoby, 2013).

Other work supports fluency’s role in hindsight bias (see Roese and Vohs, 2012). In one study, researchers presented semantically-related and unrelated prime words before playing distorted versions of target words (e.g., nurse-ducter; bread-ducter; here, “ducter” refers to the word, “doctor”). On half the trials, participants heard a clear version of the target word before hearing the prime and distorted target pair (e.g., doctor-nurse-ducter). Thus, on hindsight trials, participants knew the target word before hearing the distorted version of it. On the remaining trials, participants did not hear a clear version of the target word before hearing the prime and distorted target pair. Participants estimated how many of their naïve peers could identify the target word. This procedure yielded hindsight bias in that participants provided inflated peer estimates on hindsight trials; however, prime relatedness interacted with target knowledge as might be expected, such that hindsight bias was greater on unrelated than related prime trials. The authors argued that the fluency caused by prior target knowledge was greater and harder to ignore or discount on low-fluency, unrelated-prime trials than on high-fluency, related-prime trials (Higham et al., 2017).

In another study, researchers exposed participants to obscure facts, and later asked participants to judge how many of their naïve peers would know these facts (old) and facts not presented previously in the experiment (new). Participants judged that more of their peers would know the old than the new facts. This effect occurred when the researchers exposed participants to factual questions with and without the answers. These results further show that fluency affects hindsight bias (Birch, Brosseau-Liard, Haddock, & Ghrear, 2017).

Fluency misattribution is not the only account of hindsight bias. A sense-making account of hindsight bias focuses on the role that surprise—the discrepancy between participant expectations and the correct answer—plays in hindsight bias (Pezzo, 2003, 2011). A participant in the Clear-distorted condition would, after hearing the distorted stimulus, evaluate the difference between their expectation of what the distorted stimulus should be and the correct answer. This account of hindsight bias relies heavily on the expectations the participant has of the target (distorted) stimulus, rather than the fluency of processing the stimulus.

There also are several reconstruction-based theories of hindsight bias involving anchoring and adjustment—for example, Selective Activation and Reconstructive Anchoring or Reconstruction After Feedback with Take the best (Hoffrage, Hertwig, & Gigerenzer, 2000; Pohl, Eisenhauer, & Hardt, 2003). In these theories, hindsight bias results from a biased memory search; participants first attempt to answer trivia questions, and later attempt to recall their answers either in the presence or absence of the correct answers to the question. Participants find it more difficult to recall their prior answers in the presence than in the absence of the correct answers. The problem here is that these theories are based on reconstruction of one’s prior answers, which is of tangential relevance to the design discussed here, in which participants instead make a judgment regarding the hypothetical scenario where the clear stimulus had not been presented. This problem applies not only to those theories relying on reconstruction bias, but also many others that focus on the role of memory in hindsight bias, such as automatic assimilation or the trace-strength hypothesis (see Blank, Nestler, von Collani, & Fischer, 2008). Thus, the accounts most relevant to the auditory hindsight task that we use here are the fluency-misattribution and sense-making accounts.

The Current Work

The current work seeks to demonstrate that fluency misattribution underlies auditory hindsight bias. To this end, we modulated the influence of fluency on auditory hindsight bias by using repetition priming: Items to which a participant has been frequently primed in a separate “exposure” phase should have relatively higher fluency than items that were unprimed or primed less (see Jacoby & Dallas, 1981). This higher fluency makes subsequent identification of those items easier and
faster. There were three sources of fluency in the current study: (1) pre-experimental exposure—the baseline familiarity that participants had with the items as a consequence of the words all being commonplace and familiar; (2) the exposure phase, during which we manipulated how often a participant was primed with each item: the greater the number of priming presentations, the greater the fluency from this phase; and (3) the test phase, during which we manipulated whether a clear presentation appeared prior to the distorted target word. A fluency-misattribution account of hindsight bias predicts that these three sources combine to influence a common process, and therefore might produce an interactive effect on participant judgments, as per Roberts and Sternberg (1993).

To elicit auditory hindsight bias, we manipulated whether participants heard a clear version of a distorted target word just before hearing the distorted target word. There were two critical distortion conditions, a Clear-distorted estimation condition in which both auditory hindsight bias and repetition priming could contribute to the fluency for an item and thereby influence estimates of peer success at identifying a distorted target word, and a Distorted-only estimation condition in which the principal source of fluency was priming effects from the exposure phase. Because the target words were familiar, commonplace words, fluency from pre-experimental exposure was equal in both conditions. When estimates in the Clear-distorted condition are higher (indicating greater success at identifying target words) than estimates in the Distorted-only estimation condition, participants will have demonstrated auditory hindsight bias. We also tested a group of participants in a Distorted-only identification condition in which participants identified the distorted targets instead of providing estimates of peer success. This group provided a measure of the fluency invoked by the number of priming presentations in the exposure phase, because participants should be more accurate at identifying words that are more fluent to them.

As a manipulation of repetition priming, participants experienced four different levels of exposure to prime words (number of priming presentations) in an exposure phase. In that phase, participants heard clear versions of words 0, 1, 3, or 6 times before hearing distorted versions of those words during a later test phase. As the number of priming presentations in the exposure phase for an item increases, item fluency should increase. Similarly, item fluency should increase in the Clear-distorted condition due to the clear presentation of the item at test. The clear presentation at test is the usual source of auditory hindsight bias, whereas fluency from the exposure phase is the usual source of repetition priming. In the current work, we operationalize repetition priming as the effect of prior presentations on subsequent identification and estimation performance.

### Overview of Experiments

The current work advances prior work by combining auditory hindsight bias and repetition priming, to test and confirm a fluency-misattribution account of auditory hindsight bias. We did this by examining the interaction between our two primary manipulations: (1) presence versus absence of a clear prime prior to a distorted target at test; and (2) number of priming presentations at exposure. Specifically, a hindsight-bias effect would be evident if the clear presentation in the test phase led participants to overestimate their peers’ ability to identify distorted target items, relative to when participants were not exposed to the clear presentation at test. Repetition priming would be evident if increasing the number of priming presentations in the exposure phase increased participants’ later identification of distorted target items or estimates of their peers’ ability to identify distorted target items in the test phase (see Feustel, Shiffrin, & Salasco, 1983; Jacoby & Dallas, 1981; Ostergaard, 1998).

We wondered whether these two sources of fluency (clear prime at test, number of priming presentations during exposure) operated independently and additively or instead interacted by combining into one mechanism multiplicatively. According to Roberts and Sternberg’s (1993) additive factors logic, which was based on reaction-time data, we can infer from the data pattern whether the sources of fluency operate additively or multiplicatively. If both sources of fluency operate independently – i.e., at different stages of processing – then their effects should be additive in nature. If both sources of fluency operate together via the same mechanism—at the same stage of processing—then their effects should interact multiplicatively. An interaction, if present, could be either over-additive (whereby both sources of fluency compound their effects) or under-additive (whereby there is little or no additional benefit of having both sources present — a condition that might be characterized by each source being sufficient by itself, such as one might experience by taking both aspirin and acetomitoophen for a headache).

Because we believe that both hindsight bias and repetition priming share fluency as a common process, we expected to see an interaction—specifically, an under-additive interaction dominated by the clear presentation at test (when present). By experiencing the clear presentation at test, participants would be less likely to show priming effects from the exposure phase. The clear presentation at test should heavily influence the participants’ perceptual experience when they perceive the distorted item. The clear presentation occurred almost immediately before the distorted target...
HINDSIGHT BIAS AND REPETITION PRIMING

Figure 1. Predictions for proportion estimated success in the auditory hindsight bias task according to the fluency misattribution and sense-making accounts. Participants estimate how likely a naïve observer is to correctly identify a distorted auditory presentation of a word when that presentation occurs either following a clear presentation (Clear-distorted) or in isolation (Distorted-only). The target word is presented 0, 1, 3, or 6 times in undistorted form in an earlier exposure phase of the experiment.

word, and was perfectly correlated with the identity of the distorted target word; this stands in contrast to the primes from the exposure phase, which were presented at a relatively distant time, and which were far less correlated with the specific distorted word being presented later during the test phase. We expected that the additional fluency from the clear presentation at test and the number of priming presentations from the exposure phase would influence the encoding of the distorted word; however, the clear presentation at test, in particular, would dominate participants’ perceptual experience of the word for the reasons mentioned above, to the point that there would be no added benefit from the priming presentations in the exposure phase. Thus, for the Clear-distorted condition, we expected to see an overall increase in estimates that was generally insensitive to the number of priming presentations, resulting in an under-additive interaction (see Ostergaard, 1998). The increased fluency for the distorted word from the clear presentation would result not only in it being judged as easier to identify, but being perceived as more identifiable, independent of the number of priming presentations in the exposure phase. Unlike the Clear-distorted condition’s insensitivity to number of priming presentations, we expected to see priming effects in the Distorted-only estimation condition (see Figure 1, left panel).

Contrast this with Pezzo’s (2011) sense-making account of hindsight bias, where hindsight bias would be rooted in the “surprise” felt by the participant upon being exposed to the distorted word; if all that occurred was an expectancy violation, then the exposure phase and the clear presentation at test would not contribute to a common mechanism, and thus would not show an interactive data pattern. Instead, we would see two parallel lines (see Figure 1, right panel).

Experiment 1

In Experiment 1, we sought to manipulate the amount of fluency present during the exposure phase and the test phase. There were two critical distortion conditions, a Clear-distorted condition and a Distorted-only estimation condition. In the Clear-distorted condition, fluency from pre-experimental exposures to the items and the test phase could influence estimates of peer success at identifying a distorted target. In the Distorted-only estimation condition, fluency from pre-experimental exposures to the items and the exposure phase could influence estimates of peer success at identifying a distorted target. We anticipated that participants in the Clear-distorted condition would provide higher estimates of peer identification performance than would participants in the Distorted-only estimation condition. We also tested a group of participants in a Distorted-only identification condition in which they made overt identification responses to the distorted targets instead of providing estimates of peer
success. One purpose served by this group was to provide a manipulation check that would generate evidence that varying the number of priming presentations indeed produces differences in the fluency of processing distorted targets in the test phase.

Method

Participants. One hundred twenty-one students from Kwantlen Polytechnic University participated in the study for course credit. In all conditions, participants completed the experiment in groups of one to eight. We included three distortion conditions in the design of the experiment. Rather than randomly assigning participants to conditions, we tested the three conditions sequentially. Initially, we tested the Distorted-only identification condition \((N = 62)\) as a manipulation check to confirm that varying the number of priming presentations worked as expected. We followed this with the Clear-distorted condition \((N = 23)\), which we expected to produce hindsight bias. One might simply compare the identification and Clear-distorted conditions, but these two conditions have different dependent variables—mean accuracy in the former, mean estimates for peers in the latter. Thus, we also tested the Distorted-only estimation condition \((N = 36)\), allowing us to compare mean estimates for peers in this condition to the Clear-distorted condition. Although non-random assignment can compromise the validity of the results, Experiment 2 replicates Experiment 1’s primary results by using random assignment to conditions.

Materials. We recorded audio clips of 40 common, concrete words (e.g., leaf, muffin, dress) spoken clearly in a male voice for use in the exposure phase (and in the Clear-distorted condition of the test phase as well). We then low-pass filtered the words using a custom MATLAB script to generate distorted, difficult-to-identify versions for use in the test phase (the script is available at http://lifespancognition.org/onlineresearch/ahb_stimuli.php). We divided the 40 items into four lists of 10 items, and counterbalanced assignment of these lists to the four priming-presentation conditions \((0, 1, 3, 6)\) presentations in the exposure phase) across participants.

Design. This experiment followed a 3 (distortion condition: Clear-distorted, Distorted-only estimation, Distorted-only identification) x 4 (number of priming presentations: 0, 1, 3, 6) mixed design with distortion condition as a between-subjects factor and number of priming presentations as a within-subjects factor.

Procedure. All stimuli were presented using a CD player and loudspeaker. The experiment had two phases—an exposure phase and a test phase. The exposure phase was identical for all conditions. Participants heard the clear, undistorted words, with each word being presented 0, 1, 3, or 6 times, for a total of 100 presentations in one of four random, intermixed orders. We counterbalanced across groups of participants the number of times a given word was presented. Immediately after the exposure phase, participants began the test phase.

Each trial began with a 500-ms alerting tone. For the Clear-distorted condition, the tone was followed 1 s later by a clearly spoken version of the target word (as in the exposure phase). After a 1-s pause, the distorted version of the same target played and participants estimated the percent (proportion) of their peers that would have been able to identify the word, assuming their peers had not heard the clear presentation. The Distorted-only estimation condition skipped the clearly spoken word and the 1-s pause preceding it; the tone was followed by a 1-s pause, and then the distorted target word; participants wrote down an estimate of the percent of their peers who would be able to identify the distorted word. The Distorted-only identification condition also skipped the clearly spoken word, but after the tone and the 1-s pause, participants had 8 seconds to write on a piece of paper what they thought the word was. This procedure repeated for each of the 40 words. Participants were not told anything about how the exposure and test phases related. For an example of the procedure and stimuli, see http://lifespancognition.org/onlineresearch/ahb_stimuli.php. After completing the test phase, participants completed a short demographics form and were debriefed.

Results

The dependent variable in the Clear-distorted condition and Distorted-only estimation condition was the mean estimated proportion of peers who would be expected to identify the target (hereafter called proportion of estimated success). In the Distorted-only identification condition, the dependent variable was the proportion of words that participants identified correctly. Figure 2 depicts the mean proportion for each distortion condition as a function of the number of priming presentations. We analyzed these scores using a 3 (distortion condition) x 4 (number of priming presentations) mixed factorial analysis of variance (ANOVA). We found a main effect of distortion condition, \(F(2, 118) = 4.71, MSE = .065, p = .01\), with the Clear-distorted condition having higher mean scores than the Distorted-only identification and Distorted-only estimation conditions. This effect is evidence of auditory hindsight bias. We also found a significant main effect of number of priming presentations, \(F(3, 354) = 15.28, MSE = .020, p < .001\), with scores increasing monotonically with increasing number of priming presentations. This effect is evidence of repetition priming. Finally, we found a significant interaction between distortion condition and number of
Figure 2. Mean proportion of estimated success on the estimation task and mean proportion correct on the identification task in Experiment 1. In the estimation task one group of participants heard a clear presentation of the target word prior to the distorted version (Clear-distorted) and another group heard only the distorted version (Distorted-only) when making their estimates. An additional group of participants attempted to identify the distorted target words (Identification). All groups completed a priming phase at the beginning of the experiment in which they heard 0, 1, 3, or 6 presentations of a target item in clear form. Data points are jittered on the horizontal axis to make them more distinct. Error bars are 95% within-subjects confidence intervals suitable for comparing across number of priming presentations within a group.

priming presentations, $F(6, 354) = 3.67, MSE = .020, p = .02$. We conducted one-way ANOVAs of the number of priming presentations for each of the distortion conditions and found no effect in the Clear-distorted condition ($F < 1$), but significant effects in the Distorted-only identification and Distorted-only estimation conditions ($F[3, 183] = 20.62, MSE = .029, p < .001$ for Distorted-only identification; $F[3,105] = 9.14, MSE = .012, p < .001$ for Distorted-only estimation). Comparing Figures 1 and 2, we see more support for the fluency-misattribution account than for sense-making account of auditory hindsight bias.

**Discussion**

Performance on the Distorted-only identification task indicated that participants were sensitive to the number of priming presentations. Similar performance emerged among participants in the Distorted-only estimation condition who, rather than providing identification responses themselves, estimated the ability of their peers to identify distorted test items. In contrast to these two groups, we observed relative insensitivity to number of priming presentations in the Clear-distorted condition and a strong hindsight-bias effect: Estimates of peer success were higher in the Clear-distorted condition than in the Distorted-only estimation condition and actual identification rates in the Distorted-only identification condition, particularly for items that had few or no prime presentations. The lack of an effect of number of priming presentations in the Clear-distorted condition is consistent with the idea that fluency arising from the exposure phase and from the clear presentation both contribute to the same stage of processing—namely, the encoding of the distorted word. However, the influence of the clear presentation at test is a much stronger source of fluency than are the primes from the exposure phase, resulting in the fluency from the clear presentation at test effectively overpowering any influence from the exposure phase. This supports the fluency-misattribution account but not the sense-making account of auditory hindsight bias.

**Experiment 2**

Experiment 1 showed robust hindsight bias and repetition priming and a significant interaction between these two effects. Experiment 1 left unanswered whether task experience modifies the fluency by which participants process auditory stimuli, and by extension, hindsight bias. In Experiment 2, we used a within-subjects manipulation of distortion condition to examine the influence of task experience on hindsight bias. Half the participants completed the Clear-distorted condition first, followed by the Distorted-only estimation condition. We expected that the clear presentation of the stimulus would have its full impact for these participants because they would not yet have experienced the presentation of distorted words without the assistance of a clear presentation. Therefore, these participants should produce a hindsight-bias effect similar to that found in Experiment 1. The remaining participants completed the Distorted-only estimation condition first, followed by the Clear-distorted condition. Having experienced the Distorted-only estimation condition in the first block of trials should sensitize participants to the potential contaminating influence of the clear presentation when encountered in the second block. The reason for this is that participants should be able to compare the difficulty they had when trying to identify distorted versions of items in the first block of trials with the relative ease of identification in the Clear-distorted condition in the second block of trials. This comparison should lead participants to discount the effect of the clear stimulus in the second block of trials.

Prior work has shown that participants who were aware of a perceptual prime were able to discount its
influence on their response (Huber, Shiffrin, Lyle & Ruys, 2001; Jacoby & Whitehouse, 1989). This suggests that if participants can experience an uncontaminated version of distorted words, then they should be able to discount any further influence of the clear presentation when tested with other targets (and perhaps even the exposures in the priming phase). The result of this discounting is that there should be reduced or possibly no auditory hindsight bias (no significant difference between the Clear-distorted and Distorted-only estimation conditions) in the group that experiences the Distorted-only estimation condition first. Such a result would indicate that experience with the Distorted-only estimation inoculates participants against the biasing effects of the clear presentation at test. Such a debiasing result has not been shown in auditory hindsight bias (see Roese & Vohs, 2012, for successful debiasing strategies in other hindsight-bias work).

Method

Participants. Sixty-four undergraduate psychology students from Kwantlen Polytechnic University participated in the study for course credit; there was no overlap between this pool and the pool of participants from Experiment 1. Participants completed the experiment individually or in groups of 2, with each group being randomly assigned to one of the orders: Clear-distorted (CD) → Distorted-only estimation (D) or Distorted-only estimation (D) → Clear-distorted (CD).

Materials. This experiment used the same items as Experiment 1, with the audio clips being recorded in a female voice.

Design. This experiment followed a 2 (distortion condition: Clear-distorted, Distorted-only estimation) x 2 (order: CD → D or D → CD) x 4 (number of priming presentations: 0, 1, 3, 6) mixed design with distortion condition and number of priming presentations as within-subjects factors and order as a between-subjects factor.

Procedure. The exposure phase was the same as in Experiment 1. The test phase was modified for the within-subjects design. All participants completed both the Clear-distorted condition and the Distorted-only estimation condition, with half the items in each number of priming presentations condition appearing in each of the distortion conditions. The distortion conditions appeared in separate blocks, in one of two possible orders: CD → D, or D → CD. There was no Distorted-only identification condition in this experiment.

Results

As with Experiment 1, the dependent variable was the proportion of estimated success at identifying the distorted target word. Figure 3 depicts the mean proportion for each distortion condition as a function of the number of priming presentations. We analyzed these scores using a 2 (distortion condition) x 2 (order) x 4 (number of priming presentations) mixed factorial ANOVA with order as a between-subjects factor. We found a significant main effect of distortion condition, $F(1, 62) = 20.21, \text{MSE} = .042, p < .001$, with the Clear-distorted condition having higher mean scores than the Distorted-only estimation condition (.46 vs. .38). We also found a main effect of number of priming presentations, $F(3, 186) = 3.10, \text{MSE} = .010, p < .05$, with scores increasing monotonically with number of priming presentations. Critically, the interaction between distortion condition and number of priming presentations was significant, $F(3, 186) = 4.38, \text{MSE} = .013, p < .01$. Separate ANOVAs testing the effect of number of priming presentations in each condition indicated that repetition priming was apparent in the Distorted-only estimation condition, $F(3, 186) = 6.96, \text{MSE} = .012, p < .001$, but not in the Clear-distorted condition, $F < 1$. These results replicate those of Experiment 1. Comparison of the Clear-distorted
condition in the CD → D condition and the Distorted-only estimation condition in the D → CD condition in Figure 3, which replicate Experiment 1’s CD and D conditions, respectively, clearly shows the interaction between distortion condition and number of priming presentations that we observed in Experiment 1.

The order by distortion condition interaction was also significant, $F(1, 62) = 10.68, MSE = .042, p < .005$. We investigated this interaction by computing separate ANOVAs for each order condition. When the Clear-distorted condition occurred first, we found a substantial effect of distortion condition, $F(1, 31) = 33.03, MSE = .038, p < .001$, with higher estimates in the Clear-distorted condition than in the Distorted-only estimation condition (.50 vs. .36). When the Distorted-only estimation condition occurred first, however, we found no effect of distortion condition, $F < 1$.

**Discussion**

We anticipated that initial experience with a Distorted-only estimation condition would give participants an accurate sense of how difficult the distorted-word identification task actually was. When later confronted with a situation in which their assessment of difficulty was contaminated by hearing a clear presentation of the target word just prior to hearing the distorted version of that target word, participants appeared to engage in a form of discounting (as would be predicted by Huber et al., 2001); in an effort to account for the influence of the clear presentation, they tried to reduce their ratings. Figure 3 shows that participants were not able specifically to discount the influence of the clear presentation; rather, in accord with the fluency misattribution account, their only recourse was to apply a general reduction in their estimates, regardless of the item’s repetition history. This process reflects an overall drop in the average estimates on the Clear-distorted test in the D → CD order. We suspect that the experience with a clear presentation of the target is integrated into the fluency for the item so that subsequent efforts at discounting the influence of that experience cannot selectively eliminate its impact, exclusive of the impact of the earlier priming presentations. Consequently, any effort to discount the effect of the clear presentation should manifest as an overall reduction in estimates of peer performance, independent of the number of priming presentations.

Figure 3 shows another important result: clear discounting in the D → CD order. This result reflects the lack of difference between the Distorted-only estimation condition and Clear-distorted condition (no hindsight bias). This discounting also reveals successful debiasing: Experience with the Distorted-only estimation condition served as a debiasing technique for participants once they encountered the Clear-distorted condition. Successful discounting and debiasing are consistent with the fluency-misattribution account of auditory hindsight bias.

**Experiment 3**

Experiment 2 again showed robust hindsight bias and repetition priming and a significant interaction between these two effects. However, Experiment 2 also showed that task experience modifies hindsight bias. Experiment 2 left unanswered other means by which participants can discount the clear prime presentation at test. In Experiment 3, we provided participants with the opportunity to discount the influence of the clear presentation at test by adding a condition where it was possible to experience uncontaminated encoding of the distorted word. Here, we again replicated Experiment 1, but included an additional group of participants in which we provided a clear presentation of the target item just after the distorted presentation in the test phase. In this Distorted-clear condition, participants first heard the distorted version of the target but were instructed to withhold their estimate of peer performance until after they had heard the clear presentation. If the clear presentation is being integrated into the participants’ experienced fluency for the item and they only then proceed to make their estimate, we should see similar patterns as in the Clear-distorted condition; however, if participants instead make their judgment at the time that the distorted word is presented and are able to selectively exclude the impact of a clear stimulus which follows the distorted stimulus, then their responses should be closer to the Distorted-only estimation condition (see Sohngul, Peelle, Carlyon, & Davis, 2014).

Comparing the Distorted-clear to the Clear-distorted condition permitted us to test between the fluency misattribution and sense making accounts of hindsight bias; presenting the distorted word prior to the clear version allowed participants to form an expectation about the identity of the word prior to hearing its clear version. Consequently, a difference between the Distorted-clear and Clear-distorted conditions would be consistent with the sense-making account of hindsight bias, whereas no difference between these conditions would be consistent with the fluency-misattribution account.

**Method**

**Participants.** Two hundred seven undergraduate psychology students from Kwantlen Polytechnic University participated in the study for course credit; no students who participated in Experiments 1 or 2 participated in this Experiment. Participants completed the experiment individually or in groups of 2, with each group completing one of the distortion conditions
analyzed the .001, with scores increasing monotonically with 
condition, we found a significant effect of number of 
orelative to the other two 
next condition, Distorted-clear, resembled the Clear-
distorted condition, except that in this case, the order of 
events in a trial was tone, distorted word, then clear 
word. The instructions were similar for the Clear-
distorted and Distorted-clear conditions, in which 
participants estimated what percent of their peers would 
have been able to identify the distorted word, assuming 
they had not heard the clear word. Participants in the 
Distorted-clear condition were instructed to provide 
their estimate only after the clear presentation on each 
trial had occurred.

Results

Figure 4 depicts the mean proportion for each 
distortion condition as a function of the number of 
priming presentations. A preliminary analysis using a 4 
distortion condition x 4 (priming presentations) mixed 
factorial ANOVA on participant estimates revealed 
effects driven primarily by the vastly lower mean scores 
for the Identification condition relative to the other two 
conditions. Consequently, we analyzed the 
Identification condition on its own, and the other three 
conditions together. In a one-way ANOVA conducted 
by percent correct responses in the Identification 
condition, we found a significant effect of number of 
priming presentations, F(3,66) = 13.67, MSE = .012, p < .001, with scores increasing monotonically with increasing number of priming presentations. We then 
analyzed the estimation condition scores using a 3 
(distortion condition) x 4 (number of priming 
presentations) mixed factorial ANOVA on participant 
estimates. We found no main effect of distortion 
condition, F(2, 181) = 1.35, MSE = .094, p = .26. We 
did find a main effect of number of priming 
presentations, F(3, 543) = 8.72, MSE = .008, p < .001, 
with scores again increasing monotonically with 
increasing number of priming presentations. Most 
importantly, we found a significant interaction between 
distortion condition and number of priming 
presentations, F(6, 543) = 3.13, MSE = .008, p < .005. 
We followed up by conducting one-way ANOVAs to 
test for repetition priming in each of the four distortion 
conditions. Consistent with previous results, the 
Distorted-only estimation and Distorted-only 
identification conditions showed repetition priming 
(F(3, 240) = 18.72, MSE = .008, p < .001 for Distorted-
donely estimation; F[3, 66] = 13.67, MSE = .012, p < .001 
for Identification). Conversely, the two hindsight 
conditions did not show repetition priming (F[3, 234] = 
1.59, p > .10 for Clear-distorted; F < 1 for Distorted-
clear).

We conducted two follow-up post-hoc analyses on pairs of conditions, one comparing Clear-distorted and Distorted-only estimation (which maps directly onto the results of Experiments 1 and 2), and another with Clear-distorted and Distorted-clear (to test the prediction from the fluency misattribution account regarding discounting the influence of fluency).

Comparing the Clear-distorted and Distorted-only estimation conditions, we found no main effect of distortion condition, $F < 1$, a main effect of number of priming presentations, $F(3, 474) = 14.00, MSE = .008$, $p < .001$, and a significant interaction, $F(3, 474) = 5.94, MSE = .008, p = .001$, supporting the conclusion that repetition priming was evident only in the Distorted-only estimation condition. Comparing the Clear-distorted and Distorted-clear conditions, we found no significant effects ($F[1, 101] = 2.50, MSE = .096, p = .117$ for distortion condition; $F[3, 303] = 1.67, MSE = .008, p = .174$ for number of priming presentations; $F < 1$ for the interaction).

Discussion

As in Experiments 1 and 2, we observed that the Clear-distorted condition was insensitive to the number of priming presentations in the exposure phase, whereas the Distorted-only estimation condition was sensitive to that factor. This result is once again consistent with the fluency-misattribution account of auditory hindsight bias: The exposure phase and the clear presentation at test both contribute to a common stage of processing, with a greater contribution of the temporally near, clear presentation relative to the influence of priming presentations in the exposure phase.

One will note that for unprimed items—those items that participants did not hear in the exposure phase of the experiment—the Distorted-only identification rate was lower in Experiment 3 ($M = .10$) than in Experiment 1 ($M = .43$). This was because the stimuli used in these experiments were different; the stimuli for Experiment 1 were recorded in a male voice, whereas the stimuli for Experiment 3 were recorded in a female voice; there were also minor differences in the level of distortion of distorted items between the experiments, resulting in the items being harder to identify in Experiment 3 than in Experiment 1.

The Clear-distorted condition and its comparison to the Distorted-clear condition provided some insight into the constraints on attempts to discount the influence of the clear presentation. It appears that participants in the Distorted-clear condition did not consistently generate estimates to discount the influence of the clear presentation. If they had done so, their performance should have resembled estimates provided by participants in the Distorted-only estimation condition. Rather, participants in the Distorted-clear condition followed instructions by waiting to make their judgment until after the clear presentation was provided. The impact of the clear presentation was similar regardless of whether the clear or the distorted version of the target occurred first. As we noted in the introduction to this experiment, this outcome is consistent with the fluency-misattribution account but not the sense-making account of auditory hindsight bias.

General Discussion

Our three experiments establish a pattern of results consistent with a fluency-misattribution account of auditory hindsight bias. We presented clear versions of target words in an exposure phase 0, 1, 3, or 6 times. These items later appeared as distorted target words in a test phase. In the test phase, participants either heard a clear word they may have heard in the exposure phase, followed by a distorted word (Clear-distorted), or they heard only the distorted version of the word (Distorted-only). Participants in the test phase attempted to either identify each distorted target word or indicate how many out of 100 of their peers would be able to identify each of the distorted target words.

We observed repetition priming—from the manipulation of the number of priming presentations during the exposure phase—only in the Distorted-only conditions in which participants judged distorted targets in the absence of a clear version of the target word. Specifically, more prime presentations increased participants’ identification success and estimates of peers’ ability to identify the distorted target words. When clear versions of target words preceded (barn-bah) or succeeded distorted versions of target words (bahn-barn), the number of priming presentations during the exposure phase did not affect participants’ estimates of peers’ ability to identify the distorted target words. Moreover, estimates of peers’ ability to identify the distorted target word were higher when a clear version appeared before the target, relative to when no clear presentation appeared, constituting a robust hindsight-bias effect (see also Bernstein et al., 2012; Harley et al., 2004).

This insensitivity to the number of priming presentations induced by the clear presentation at test is consistent with the fluency-misattribution account. In terms of Roberts and Sternberg’s (1993) additive factors method, both the repeated primes in the exposure phase and the clear presentation at test contribute to the fluency with which participants process a given item; if these sources of fluency were operating independently—if they were involved in different stages of processing—we would expect to see an additive pattern of effects. Instead, we see an interactive data pattern, suggesting that both sources funnel into a single common mechanism (namely,
fluen
cy). The insensitivity to the number of priming
presentations—an under-additive interaction with
the influence of the clear presentation at test—is a
consequence of the clear presentation at test (1)
occuring much closer in time to the actual judgment
than the repeated primes during the exposure phase, and
(2) being known to be related to the identity of the
distorted word (in contrast to items in the exposure
phase, which appeared in one essentially homogenous
block prior to the test phase). Overall, our data support
the fluency-misattribution account, rather than the
sense-making account, of auditory hindsight bias.

A possible concern when claiming evidence for a
fluency-misattribution account is that the current study,
like many others in the literature, lacked an independent
measure of fluency. We have, however, our
manipulation check: the Distorted-only identification
condition; the repetition-priming effect observed in this
condition is evidence that increased number of prime
presentations increased fluency. In our experiments, the
Distorted-only identification conditions showed a
pattern of results parallel to that found in the Distorted-
only estimation condition. Although the Distorted-only
identification condition was conducted on an
independent group in both cases, we can reasonably
infer that this condition is a viable proxy for an
independent measure of fluency. Future studies should
provide an independent measure of fluency on each
trial, such as reading time, that can then be correlated
with peer estimates (see Undorf, Zimdahl, & Bernstein,
2017).

Links to Hindsight Bias

Our hindsight bias results differ from those in the
literature. Specifically, we found that prior
experimental exposure to clear versions of words
(number of priming presentations) increased
identification success (repetition priming); however,
number of priming presentations had no noticeable
effect on hindsight bias. Previous work has shown that
compared to new trivia, prior exposure to trivia increased participants’ thinking they knew the
trivia all along (Wood, 1978) or estimates of peers’
ability to know the trivia (Birch et al., 2017; see also
Harley et al., 2004). We believe that the different data
patterns reflect stimulus and procedural differences.
Birch et al. and Wood used trivia questions, and Harley
et al. used celebrity faces.

Other work involving trivia shows that the more
available the correct answer is when trying to recall
one’s original, naïve answer or estimating what a naïve
peer would answer, the more hindsight bias results
(Groß & Bayen, 2015; Hell, Gigerenzer, Gauggel, Mall,
& Muller, 1988). It is likely that Wood’s (1978) and
Birch et al.’s (2017) participants found it hard to ignore
the correct answers to trivia that they had studied
multiple times, resulting in anchoring on the correct
answer and insufficiently adjusting from it when
estimating from a naïve perspective (see Epley, Keysar,
Van Boven, & Gilovich, 2004). Our participants faced a
somewhat different problem: They encountered clear
versions of target words in an exposure phase, and later
during a test phase, received an additional exposure to
tothose clear words immediately prior to hearing a
distorted version of the words. The test-phase exposure
to clear words appears to have contributed substantially
to the participants' fluency for the word, overpowering
the influence of priming-phase exposure to those words.

Harley et al.’s (2004, Experiment 3) stimuli and
procedure also differed from those in the current work.
During a Distorted-only phase, Harley et al. had
participants identify celebrity faces as those faces
gradually clarified from full blur to full clarity. Later,
during a Clear-distorted phase, participants saw clear
versions of celebrity faces before those faces again
clarified from full blur to full clarity. Half the test faces
had been seen in the Distorted-only phase (old celebrity
faces), and half had not (new celebrity faces).
Participants indicated when a naïve peer would identify
the celebrities, had their peer not seen the clear version
of the celebrity face. Harley et al. found robust
hindsight bias for both old and new celebrity faces in
that participants stopped the clarification of faces
sooner during the Clear-distorted phase than during the
Distorted-only phase. Critically, though, participants
showed more hindsight bias for old compared to new
celebrity faces.

Harley et al. (2004) interpreted these results in
terms of fluency misattribution; participants process
distorted versions of celebrity faces more fluently when
they have encountered both distorted and clear versions
of those faces previously in the experiment. Because
participants identified gradually-clarifying celebrity
faces during both Distorted-only and Clear-distorted
phases, participants experienced massive fluency for
old celebrity faces. Specifically, participants saw 30
levels of blur as each celebrity face clarified during the
Distorted-only phase, and the identical 30 levels of blur
as each old celebrity face clarified during the Clear-
distorted phase. This massive fluency likely contributed
to the increased hindsight bias for old versus new
celebrity faces, because participants had trouble
discounting fluency’s effects of their ability to identify
blurred celebrity faces. Notably, Harley et al. compared
no prior presentation to 30 prior presentations of a
celebrity face. Using a within-subject design, we
compared no prior presentation in the Distorted-only
estimation condition to several different levels of prior
presentation (up to a maximum of 6) in the Clear-
distorted condition. Perhaps had we compared no prior
presentation to 6 prior presentations only, we might
have observed a difference between our 0- versus 6-
presentation conditions (visual inspection of Figures 2-
4 shows no difference between 0 versus 6 presentations). Hence, it is possible that distinctiveness accounts for differences between ours and Harley et al.’s results: In our experiments, old items were indistinct from new items. In Harley et al.’s experiment, old items were distinct from new items.

In light of our fluency-based account, the following picture emerges: Hindsight bias arises through a fluency-misattribution process in which participants fail to discount or ignore knowledge acquired through prior exposure to that knowledge, with number of priming presentations contributing heavily to that fluency. Why, then, do we see no benefit of number of priming presentations for hindsight items? As noted earlier, the single clear exposure at test is strong enough to overpower the combined effect of the clear presentations during the exposure phase; participants’ estimates seem to be determined primarily by exposure to the clear stimulus in the test phase, obscuring the effect of number of priming presentations from the exposure phase. That this does not happen in Harley et al. (2004) may be because in their work, the clear presentation at test (hindsight condition) was not strong enough to overpower the 30 levels of blur or priming which the participant had seen previously for old faces. Rather, Harley et al.’s participants likely encoded the clear presentation at test separately from the 30 levels of priming.

Another novel aspect of our findings involves the debiasing effect that we observed in Experiment 2 in the Distorted-only estimation first condition. Here, participants experienced the difficulty firsthand of hearing distorted words and trying to estimate how many of their peers could identify those words. This initial experience with the Distorted-only estimation condition gave participants an accurate sense of how hard the distorted-word identification task was. When they later encountered a situation in which their former knowledge was acquired through prior exposure to a distorted stimulus, we think that we and others knew it, saw it, or heard it all along. Such hindsight bias can hamper our ability to appreciate our own prior or other people’s struggles with uncertain judgments. Sometimes hindsight bias is benign; other times it is harmful, such as in the case of the Italian court finding a group of scientists and a former government official guilty of failing to predict a catastrophic earthquake. Whether benign or harmful, hindsight bias makes an uncertain world more certain.

References

Fischhoff, B. (1975). Hindsight = foresight: The effect of


