AMERICAN PSYCHOLOGICAL ASSOCIATION

Journal of Experimental Psychology: General

ISSN: 0096-3445

2023, Vol. 152, No. 2, 464–482
https://doi.org/10.1037/xge0001284

Curiosity: The Effects of Feedback and Confidence on the Desire to Know

Janet Metcalfe[1], Matti Vuorre[2], Emily Towner[3], and Teal S. Eich[4]
[1] Department of Psychology, Columbia University
[2] Oxford Internet Institute, University of Oxford
[3] Department of Psychology, University of Cambridge
[4] Leonard Davis School of Gerontology, University of Southern California

In 10 experiments, we investigated the relations among curiosity and people's confidence in their answers to general information questions after receiving different kinds of feedback: yes/no feedback, true or false informational feedback under uncertainty, or no feedback. The results showed that when people had given a correct answer, yes/no feedback resulted in a near complete loss of curiosity. Upon learning they had made an error via yes/no feedback, curiosity increased, especially for high-confidence errors. When people were given true feedback under uncertainty (they were given the correct answer but were not told that it was correct), curiosity increased for high-confidence errors but was unchanged for correct responses. In contrast, when people were given false feedback under uncertainty, curiosity increased for high-confidence correct responses but was unchanged for errors. These results, taken as a whole, are consistent with the region of proximal learning model which proposes that while curiosity is minimal when people are completely certain that they know the answer, it is maximal when people believe that they almost know. Manipulations that drew participants toward this region of "almost knowing" resulted in increased curiosity. A serendipitous result was the finding (replicated four times in this study) that when no feedback was given, people were more curious about high-confidence errors than they were about equally high-confidence correct answers. It was as if they had some knowledge, tapped selectively by their feelings of curiosity, that there was something special (and possibly amiss) about high-confidence errors.

Keywords: curiosity, confidence, conflict monitoring, metacognition, region of proximal learning

Supplemental materials: https://doi.org/10.1037/xge0001284.supp

This article investigates the impact of feedback upon epistemic curiosity (Berlyne, 1954; Loewenstein, 1994) or what is also called "Curiosity1" (Metcalfe & Jacobs, in press)[1]—the goal-directed desire to know the correct answer to a specific question. Feedback can provide yes/no information regarding the person's accuracy: Is the answer they provided correct or incorrect? Alternatively, feedback can be informational and assert an answer. The provision of an answer, though, does not definitively imply that the answer that is so provided is true and correct. In this series of 10 interrelated experiments, we investigated people's curiosity in relation to their confidence in their own answers under several variants of feedback, including the provision of no feedback at all. Our objective was to manipulate people's certainty that they knew the answer, and to chart the effects on their curiosity.

Our focus is on the predictions of the region of proximal learning (RPL) framework of curiosity (Metcalfe et al., 2020), though we make reference to other models as well. The basic tenant of the RPL view is that people are most curious—they are most eager to

This article was published Online First September 1, 2022.
Janet Metcalfe https://orcid.org/0000-0003-3286-9475
This research was supported by National Science Foundation Grant 1824193 awarded to Janet Metcalfe. The authors are solely responsible for the content of this article. This study was not preregistered. The supporting data and code are available on the Open Science Framework (https://osf.io/akxs4/).

Parts of the results reported in this article were presented at the 59th Annual Meeting of the Psychonomic Society, November 17, 2018, New Orleans, Louisiana, United States; at the 61st Annual Meeting of the Psychonomic Society, November 21, 2020, Virtual; at Sarmac XIII, June 8, 2019, Cape Cod, Massachusetts, United States; and at a UCLA colloquium, October 9, 2020, Virtual.

Correspondence concerning this article should be addressed to Janet Metcalfe, Department of Psychology, Columbia University, 401B Schermerhorn Hall, 1190 Amsterdam Avenue, New York, NY 10027, United States. Email: jm348@columbia.edu

[1] Note that Metcalfe and Jacobs (in press) argued that there is a second kind curiosity, which they call Curiosity2. Unlike Curiosity1, it is goal neutral or possibly goal averse, and is characterized by exploration and play, rather than by the desire to get the rewarding answer. They argued that this discursive kind of curiosity is often confounded with Curiosity1. Jacobs and Metcalfe further proposed that Curiosity1 and Curiosity2 are anchored in different neural systems. According to this view, the habit-related, reward/goal-oriented, reinforcement learning based (RL), striatally-centered S-R neural system is associated with Curiosity1, while the reward neutral, "novelty-philic," episodic memory, default-mode, incidental learning, cognitive, hippocampally-centered, S-S neural system is associated with Curiosity2.

know and to learn—when they feel they are close to knowing the answer. Many other theorists have also proposed that "almost knowing" piques people's curiosity and primes a hunger to know (Berlyne, 1954; Kang et al., 2009; Loewenstein, 1994; Marvin & Shohamy, 2016; Metcalfe et al., 2017, 2020; Murayama, 2021). Berlyne (1954) noted that items that are neither too easy nor too difficult are those about which people are most curious. Atkinson (1972), in his seminal model of learning, called such items *transition items* and demonstrated that selective machine-guided study of them resulted in enhanced learning. Such items are thought to be in an individual's own RPL (Kornell & Metcalfe, 2006b; Metcalfe et al., 2020). Metcalfe and Kornell (2003, 2005) and Xu and Metcalfe (2016) have shown that when asked to choose what they want to study, people select materials in their own RPL, and doing so results in improved learning. High feeling of knowing items show similar results (Brooks et al., 2021).

Tip-of-the-tongue (TOT) items are classic examples of Curiosity1-evoking items that are on the verge of being known (Schwartz, 1999), and, as might be expected, TOT items have been shown to be associated with very high levels of curiosity (Metcalfe et al., 2017). On receipt of the correct response to TOT items, people's encoding, as evidenced by a distinctive ERP signature, and their later recall to such items, has been shown to be enhanced (Bloom et al., 2018).

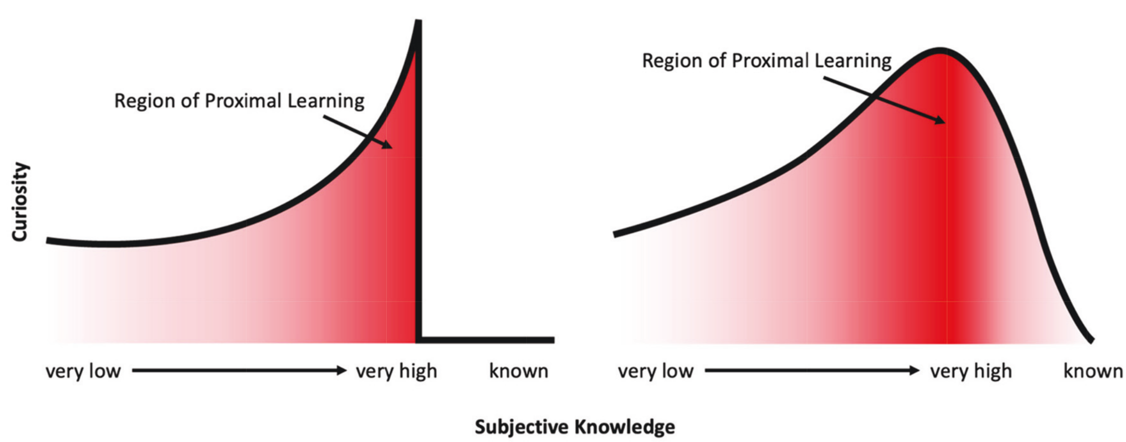
Curiosity is thought to depend on metacognition (FitzGibbon et al., 2020; Kang et al., 2009; Litman, 2005; Litman et al., 2005; Loewenstein, 1994; Metcalfe et al., 2020). If people believe that they are far from knowing the answer or if they are convinced that they already know the answer, they should not be curious. Thus, metaknowledge that they almost know may be critical and necessary for curiosity to be aroused. Indeed, as noted by Loewenstein (2007): "A failure to appreciate what one does not know would constitute an absolute barrier to curiosity" (p. 161).

Berlyne (1954) noted that once an answer is known, the drive-to-know (curiosity) is immediately quenched. If the person knows for sure that they have the correct answer, they should not be curious. Definitive feedback that one is correct, then, should satisfy curiosity. On the other hand, being told that they were wrong when they were certain that they were right should result in a great deal of curiosity.

An illustration of the RPL model (which we have modified from Metcalfe et al., 2017), showing how curiosity would be captured in that framework, is presented in Figure 1. In the left-hand panel there is a clear threshold between items that are known and not known. If the person is correct and is given definitive feedback that they are correct, they should be to the right of the threshold indicating that they "know" the answer. As such, they should show no curiosity. If, however, they receive yes/no feedback and learn that they are incorrect, the RPL model predicts that the greater was their confidence in their answer, the closer they should be to the threshold of knowing, and the greater their curiosity. On the other hand, if people are not provided with definitive feedback about their knowledge, then the boundary between known-items and high-confidence items is likely to be fuzzy, as is shown in the right panel of Figure 1, and curiosity may vary depending on small changes in confidence, especially among high-confidence items. We explore both aspects of this model.

In the current series of experiments, we investigated the relation between people's curiosity and their confidence about whether they knew the answer, using feedback manipulations that altered their certainty or their belief that they knew the answer. The first set of experiments, Experiments 1 through 4, investigated the effects of yes/no feedback as compared to receiving no feedback. In these, we looked at effects of feedback that provide a clear knowledge boundary. The second set of experiments, Experiments 5 through 7, investigated the effects of correct-answer (i.e., true) feedback, but under conditions of uncertainty: that is, when the correct answer was given but participants were not told that it was correct. The boundary between what was known and not known, in these cases, was fuzzy. The third set of experiments, Experiments 8 through 10, again examined a situation in which the boundary was fuzzy, but this time investigating the effects of false but uncertain feedback on curiosity.

Figure 1
The Region of Proximal Learning Model



Note. The region of proximal learning model with a hard boundary (left panel) indicating that the person has knowledge of when an item is objectively "known," such as occurs when yes/no feedback is provided and a soft boundary (right panel) for when the distinction between known and unknown is noisy, that is, when the truth of the feedback is uncertain. See the online article for the color version of this figure.

Part A. The Effect of Yes/No Feedback on Curiosity

Experiment 1: Effects of Yes/No Feedback on Curiosity About Errors

In the first experiment, we put an emphasis on errors, insofar as the errors that an individual commits constitute a particular opportunity for learning (Hays et al., 2013; Metcalfe, 2017). We were especially interested in the relation between high-confidence errors and curiosity. The question addressed was: are high-confidence errors like known answers which evoke little curiosity, or do people treat high-confidence errors as "almost" but not quite known and, hence, like TOT items that evoke a great deal of curiosity?

High-confidence errors seem similar to TOT items in many ways. For instance, correct answers to high-confidence errors are hyper-encoded and recalled better than are correct answers given as feedback to low-confidence errors (Butterfield & Metcalfe, 2001, 2006; Cyr & Anderson, 2013; Eich et al., 2013; Fazio & Marsh, 2010; Metcalfe et al., 2012; Metcalfe & Finn, 2012; Sitzman et al., 2020; Van Loon et al., 2015), just as answers to TOTs are hyper-encoded and remembered better than answers to non-TOTs (Bloom et al., 2018; Schwartz, 1999). When a person is in a TOT state, they have considerable information about the answer—the first letter, the number of syllables, semantic information (Brown, 1991; Kornell & Metcalfe, 2006a). Similarly, when people make high-confidence errors they appear to have considerable information: they make accurate second guesses; they can produce the correct answer with only a few cues; they tend to choose the correct answer on a multiple-choice test (Metcalfe & Finn, 2011, 2012). Furthermore, the high-confidence erroneous responses are close to the correct answers in semantic space (Metcalfe & Finn, 2011; Sitzman et al., 2014, 2015); just as are blockers in TOT states (Kornell & Metcalfe, 2006a). Finally, the corrections to high-confidence errors elicit a distinctive event-related potential (ERP) associated with encoding (Butterfield & Mangels, 2003; Metcalfe et al., 2015) that is similar to that observed to the resolution of TOT states (Bloom et al., 2018).

But, despite these similarities, when people make high-confidence errors they metacognitively indicate that they are highly confident that they have the answer already, whereas when they are in a TOT state, they know that they do not have the answer. Given that people make their confidence ratings sincerely, curiosity about the answers to high-confidence errors should be expected only if people are given external feedback that they were wrong.

We have been able to find only two studies that investigated curiosity and errors, but neither allowed us to definitively address the relation between high-confidence errors and curiosity. In the first study, Wade and Kidd (2019) showed that confidence in errors was weakly, but positively, related to curiosity. Unfortunately, they eliminated 13% of their data—the critical questions that people answered incorrectly but thought they had gotten right (i.e., the high-confidence errors). In the second study, Fastrich et al. (2018) asked people to give answers to curiosity-evoking questions and then to make confidence judgments about the correctness of their responses. Only the errors were examined. A positive relation between confidence and curiosity was found. However, in this study, curiosity ratings were made before people were given the feedback that they were wrong. Participants, presumably, still believed—for the high-confidence errors, at least—that they were correct, and, as such, it is puzzling that the relation between confidence and curiosity was positive. Curiosity for high-confidence correct responses was not reported. Thus, the results of this study escape easy interpretation.

In our first experiment, when people made an error, we told them they were wrong (without giving them the correct answer), and only then asked for curiosity ratings. The hypothesis was that high-confidence errors would be firmly in the learners' RPL and the pattern of curiosity should mirror that shown in the left panel of Figure 1, with curiosity monotonically increasing with confidence in the error.

Method

Participants. Forty-four (34 female, nine male, one did not answer) Columbia University and Barnard College students (age = 18–44 years) completed the study. The number of participants, here and in all experiments in this series, was based on many previous experiments from our laboratory (e.g., Metcalfe et al., 2017, 2021) which produced highly reliable curiosity-related results in similar paradigms with between 25 and 45 participants. They received course credit for participating in the experiment. The procedures in all experiments were approved by the Columbia University Institutional Review Board (Protocol No.: AAAD4902).

Procedure. Ninety general information questions (see the online supplemental material) from Nelson and Narens' (1980) norms, as updated in Bloom et al. (2018), were presented in a random order. The participant typed in an answer and then made a confidence judgment about the correctness of their answer on a sliding scale from 0% (*not at all confident*) to 100% (*completely confident*). They were then given yes/no feedback about the correctness of their answer. If incorrect, they rated their curiosity to find out the correct answer on a sliding scale from 0% (*do not care*) to 100% (*care very much*).

Results

The data were hand-checked to validate the automatic labeling of correct and incorrect answers, resulting in the elimination of 87 observations (for example, the participant used two words instead of one [e.g., "the Titanic" as compared to "Titanic"]). Participants answered an average of 34% (SE = 2%) of questions correctly, leaving 2,568 error trials where both confidence and curiosity were provided.

We analyzed these data with a multilevel regression model, where curiosity was regressed on linear and quadratic terms of confidence. We treated all effects as random across participants and included random intercepts for items. Although model comparison via Watanabe–Akaike information criterion (WAIC) did not favor including the quadratic term of confidence in this first experiment, we included it for consistency with subsequent experiments, where model comparison favored its inclusion (see below). We summarize the model's parameters with the posterior 95% quantiles (credibility interval [CI]), and the posterior probability of direction p_d, which indicates the probability that the effect is in the observed direction and is numerically equivalent to one minus the one-sided frequentist p-value. For convenience, we label parameters whose p_d is greater than 97.5% as "significant." We estimated the model with the brms R package (Bürkner, 2017; R Core Team, 2019).

Confidence in the correctness of errors and participants' curiosity were positively associated, as predicted; the quadratic component was not significant (b_{Linear} = .24, *SE* = .04, 95% CI [.16, .32], $p_d > 99.9\%$; $b_{\text{Quadratic}}$ = .02, *SE* = .03, 95% CI [−.03, .07], p_d = 78.4%; see Figure 2).

Discussion

Experiment 1 indicated that there was a positive relation between confidence in one's errors and curiosity, upon receiving feedback that one's answer was incorrect. The pattern of results is similar to the expectation of what should happen in the RPL model, as shown in Figure 1. This result is not puzzling. People had just found out that they were wrong when they thought they had been right. But they were objectively (according to past research) very close to having the answer—exactly the conditions that most researchers agree should provoke curiosity. They appreciated that they were wrong, presumably because of the feedback so indicating, but they felt close to having the answer.

But why did Fastrich et al.'s (2018) participants not show a steep decline in curiosity at very high levels of confidence, given that, when they made their curiosity judgments, they believed that their high-confidence errors were correct? To further investigate this issue, and the potential difference between their experiment and ours, Experiment 2 manipulated whether people were or were not told they were wrong, before asking for their curiosity ratings. The results for questions for which the answers were correct were also analyzed as a function of confidence, allowing us to explicitly check that when people were certain that they knew the answer—because they were explicitly told that they were correct—their curiosity was, indeed, quenched.

Figure 2
Experiment 1: Curiosity Following Yes/No Feedback As a Function of Confidence in Errors

Note. Thin lines indicate participant-specific regression lines; the bold line and shaded ribbon indicate the average regression line and its 95% credibility interval. See the online article for the color version of this figure.

Experiment 2: Effect of Yes/No Feedback as Compared With No Feedback on Correct Answers and Errors

Experiment 2 investigated curiosity for both correct and incorrect answers when yes/no feedback was provided following people's answers, as well as when no feedback was given. We expected to replicate the results of Experiment 1 in the incorrect answer/feedback condition. The RPL model predicts that without feedback, though, the boundary between what is known and what is not known will be noisy. As a result, curiosity should peak at a high level of confidence, but then should decline (i.e., it should be nonmonotonic) as is illustrated in the right panel of Figure 1. In the absence of feedback, it was predicted that curiosity for high-confidence correct and incorrect responses should be the same because people presumably have no way of distinguishing between their own correct and incorrect responses, other than what is reflected in their confidence ratings. Finally, as all theories, including RPL, predict, if people were correct and received definitive feedback that they were right, as is the case with yes/no feedback, their curiosity should be sated and fall close to zero, because the answer would then be known to them with certainty.

Method

Eighty (42 female, 36 male, two did not respond) Columbia University and Barnard College students (age range = 18–30 years) completed the study. A larger sample was used here than was used in Experiment 1, because this experiment included both a feedback and no-feedback condition (within subjects). Participants received course credit for participating. The stimuli were 128 general information questions. The procedure was similar to that of Experiment 1 except that on a random half of the trials, participants were given yes/no feedback about the correctness of their answer, whereas on the other half of trials, they got no feedback. Once participants received (or did not receive, in the no-feedback condition) feedback, they rated their curiosity to find out the correct answer.

A learning phase and a testing phase then followed, using only the no-feedback items from the initial phase. There were two groups: The first group studied (for 4 s each) the questions and answers associated with their lowest curiosity ratings (see the online supplemental material), whereas the other group studied the items associated with their highest curiosity ratings. After studying all items, they were tested. The experiment was programmed with *jsPsych* (de Leeuw, 2015).

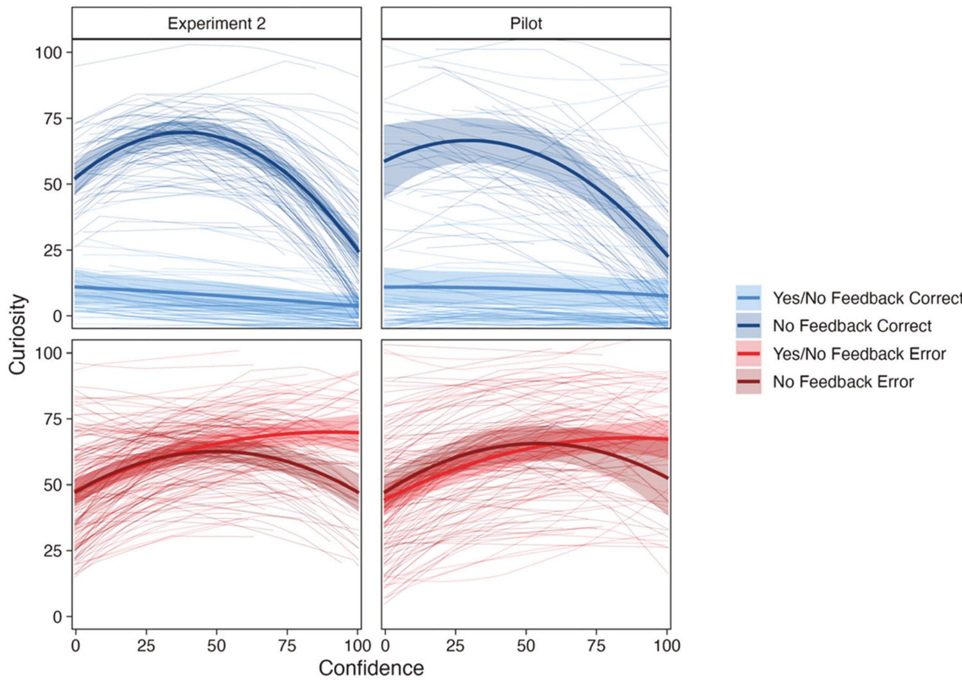
An otherwise identical pilot experiment (*N* = 72, age = 18–37 years) was conducted, but because of a programming error, the number of trials in the feedback conditions was not balanced (feedback was provided in 88% of trials). Nevertheless, this pilot experiment was very similar to Experiment 2 and provides evidence concerning replicability. The results of the pilot experiment are shown in Figure 3 along with those of Experiment 2.

Results

Effect of Feedback. Fifty-four observations were removed due to incorrect determination of accuracy (such as trivial spelling errors) by the program, leaving a total of 10,186 trials for analysis. The overall percent correct across participants was 24% (*SE* = 1%). We modeled the data in a manner that was similar to that of

Figure 3

Experiment 2 and Pilot Experiment: Curiosity for Correct Answers and Errors



Note. Curiosity for correct answers (top panels) and errors (bottom panels) depending upon whether yes/no feedback was given or feedback was not provided (no-feedback condition). Left panel: Results of Experiment 2. Right panel: Results of the pilot experiment. See the online article for the color version of this figure.

Experiment 1, but also included feedback condition and response accuracy, and all their interactions, as fixed and random effects over participants. Model comparison using WAIC strongly favored including the quadratic term of confidence in the model (WAIC difference = 148, $SE = 20$).

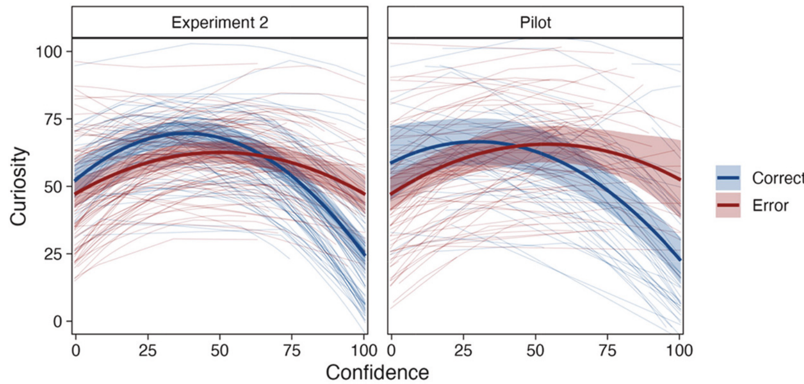
The effect of feedback was dramatic for items on which the person had been correct, as can be seen from the top panels of Figure 3. For their correct responses when there was no feedback provided, participants were quite curious to know the answer and there was a peak in curiosity at moderate levels of certainty and a decrease in curiosity as their certainty that they had been correct increased ($b_{\text{Intercept}} = 52.98$, $SE = 2.27$, 95% CI [48.62, 57.59], $p_d > 99.9\%$; $b_{\text{Linear}} = -.27$, $SE = .04$, 95% CI [−.35, −.20], $p_d >$ 99.9%; $b_{\text{Quadratic}} = -.31$, $SE = .03$, 95% CI [−.36, −.25], $p_d >$ 99.9%). This pattern is what would be expected by the RPL model when the criterion is uncertain, as in the right panel of Figure 1. Once they were told that they were correct, however, curiosity fell to close to zero, just as was expected ($b_{\text{Intercept}} = 8.97$, $SE = 2.50$, 95% CI [4.20, 13.94], $p_d = 99.9\%$; $b_{\text{Linear}} = -.07$, $SE = .03$, 95% CI [−.14, −.01], $p_d = 98.8\%$; $b_{\text{Quadratic}} = .00$, $SE = .03$, 95% CI [−.06, .05], $p_d = 56.4\%$).

Erroneous responses are shown in the bottom panels of Figure 3. When participants made an error and were told that they had made an error, curiosity at the low end of confidence stayed about the same as was the case without feedback (difference at 10: −.32, $SE = .93$, 95% CI [−2.23, 1.41], $p_d = 62.7\%$): they knew they were guessing so the feedback that they had been wrong did little to change their curiosity. However, curiosity increased selectively and significantly for errors that had been made with high confidence (difference at 90: 17.28, $SE = 3.16$, 95% CI [11.14, 23.5], $p_d = 100.0\%$).

Curiosity for Correct Answers and for Errors When No Feedback Was Provided. The data presented were rearranged to illustrate a serendipitous finding (see Figure 4): When no feedback was provided, there was a difference in participants' curiosity depending upon whether they had answered correctly or incorrectly. There was no overall linear effect of confidence on curiosity for errors though the quadratic effect was significant ($b_{\text{Linear}} = .00$, $SE = .04$, 95% CI [−.08, .07], $p_d = 50.6\%$; $b_{\text{Quadratic}} = -.16$, $SE = .03$, 95% CI [−.21, −.11], $p_d > 99.9\%$); whereas for correct answers, both the linear and quadratic terms were significant, with the linear effect being negative ($b_{\text{Linear}} = -.27$, $SE = .04$, 95% CI [−.35, .20], $p_d > 99.9\%$; $b_{\text{Quadratic}} = -.31$, $SE = .03$, 95% CI [−.36, −.25], $p_d > 99.9\%$). Furthermore, the difference in the linear confidence-curiosity association between errors and correct answers, when no feedback was presented, was significant (.27, $SE = .04$, 95% CI [.19, .36], $p_d > 99.9\%$). Similarly, the difference between errors and correct responses (when no feedback was given) was significant in the quadratic relations (.15, $SE = .03$, 95% CI [.08, .22], $p_d > 99.9\%$). To further examine this unexpected difference in curiosity as a function of response accuracy we conducted a post hoc test comparing high confidence (operationalized as confidence = 90) errors and correct answers when feedback was not provided. This contrast indicated that high-confidence errors were associated with greater curiosity than were high-confidence correct answers (Difference = 14.69, $SE = 2.56$, 95% CI [9.86, 19.91], $p_d > 99.9\%$).

Figure 4

Experiment 2 and Pilot Experiment: Curiosity Ratings As a Function of Confidence for Correct Answers and Errors With No Feedback



Note. See the online article for the color version of this figure.

Replication. The pilot experiment's results replicated those of Experiment 2, overall, as is shown in the right panels of Figures 3. Furthermore, as is shown in the right panel of Figure 4, the unexpected difference in curiosity for high-confidence correct answers as compared to high-confidence errors when feedback was not provided was observed in the pilot experiment. A post hoc contrast on the pilot study data of curiosity at Confidence = 90 indicated that high-confidence errors were associated with significantly greater curiosity than were high-confidence correct answers (Difference = 23.39, *SE* = 5.73, 95% CI [12.43, 34.89], $p_d > 99.9\%$).

Recall. Finally, the recall data showed that the group that studied and was tested on high-curiosity items recalled more on the posttest (M = .81, SE = .02) than did the group that studied their low-curiosity items (M = .63, SE = .03, $t[59] = 4.64$, $p <$ .001).

Discussion

These results indicate that yes/no feedback had a large effect on curiosity. For correct answers, when people were told that they were correct, their curiosity was almost completely sated. This finding is consistent with virtually all theories of curiosity: Given that curiosity is defined as "the desire to know," once the person does know, there is no further need to be curious.

Yes/no feedback that indicated to people that their high-confidence errors were wrong increased people's curiosity, selectively, for those errors. Feedback that one was wrong made little difference to people's curiosity about their low confidence errors (about which people remained only modestly curious both with and without feedback). The finding of a relative lack of curiosity about low confidence errors is consistent with the idea that people are uninterested in materials that are outside their own RPL and about which they know little.

Finally, when people were not given feedback about whether they were correct or incorrect, they were, nevertheless, more curious about high-confidence errors than they were about high-confidence correct answers. This was a surprising finding, suggesting that somehow making a judgment of curiosity might have either alerted people who claimed to be highly confident about their answers to the possibility that they might be wrong, or, perhaps, that feelings of curiosity are a marker that, at least some of the time, an answer given with high confidence as being correct might be an error.

Experiment 3: Second–Order Subjective Uncertainty About Correct and Incorrect Responses

Experiment 3 investigated the unexpected finding from Experiment 2 that people's curiosity ratings were higher for high-confidence errors than they were for equally highly confident correct responses. This finding suggests the possibility that they knew something more about the (lack of) truth of their answers than their own just-stated confidence indicated.

One potential explanation for the difference is that individuals might sometimes make judgments about the accuracy of their responses impulsively. When probed to consider further (as they might have done spontaneously when asked to make the curiosity judgment), they might have reevaluated, engaging in second-order metacognition that could potentially be different from their first-order judgments of confidence (Schwarz & Vaughn, 2002). If primarily confirmatory evidence is considered in confidence judgments, and disconfirming evidence is usually ignored but could be brought to bear if people were pressed further, then people might detect the possibility that they had made a mistake only when they were asked to make a second-order metacognitive judgment. The curiosity ratings might have served as such a second-order judgment, causing people to reflect further on the possibility of being wrong (and to be more curious, because now the answers were "almost known" rather than "known").

This possibility draws on several psychological theories concerned with how people appraise truth (Gilbert, 1991; Gilbert et al., 1990; Koriat, 1993, 2008a, 2008b, 2012; Lewandowsky et al., 2012). One possibility forwarded is that people's default belief, upon retrieval of remembered facts or upon comprehension of externally presented propositions, is not a balanced consideration of whether the "fact" is either true or false. Instead, if the proposition under consideration is easily understood (or fluent) it is assumed to be true. This epistemic account of the ascertainment of true beliefs fits well with many studies that have shown that confidence is positively related to the fluency of retrieval (e.g., Kelley & Lindsey, 1993; Kelley & Jacoby, 1996; Koriat, 2008b; Undorf, 2021; Wang et al., 2016; and see Schwarz, 2015 for a review). It is only upon further reflection that people may notice features or information that make them reconsider. Such noticing of the possibility that certain propositions may be false (i.e., the flagging of doubt) may not be automatic. Doubt flagging—a process that Asp et al. (2013) proposed is both slower and has different neural correlates than the initial judgment—might be evoked only if a second-order judgment is required.

There is a suggestion of support for the possibility that second-order metacognitive judgments can, under some circumstances, conflict with first-order judgements. Dunlosky et al. (2005) found that asking people to rate how confident they were in their first confidence rating brought those judgments into slightly closer alignment with the truth. Buratti et al. (2013) had people answer semantic memory questions and rate their confidence about their answers. Although, overall, second-order judgments were very similar to first-order judgments, when people were later asked if they wanted to change any of their confidence ratings, the few that people did change tended to be errors on which they had been confident.

Experiment 3, then, endeavored to press people to consider that they might be wrong by having them make a second-order metacognitive judgment. After making their confidence rating, participants were asked to rate how surprised they would be if they found out that the answer they had just given was wrong. The hypothesis was that the second-order metacognitive judgments of subjective uncertainty would reveal a difference between errors and correct responses similar to that found, in Experiment 2, with curiosity ratings.

Method

Because of COVID-19, we were unable to test college students in person as in Experiments 1 and 2. Experiment 3, and the remaining experiments in this series, were therefore conducted online. Forty-eight adults were recruited to the study using Prolific (www.prolific.com). Participants in this and all further experiments were between the ages of 18 and 45 and were required to be native English speakers. Participants averaged 24.4 ∓ 6.80 years. Thirty-one participants identified as female and 15 as male, one identified as "other," and one declined to answer. Sixteen participants reported that high school was their highest level of education, three reported associate's degree, 28 reported some college (with 17 of them reporting that they were currently in college), and one declined to answer. Participants in this and all subsequent experiments were paid at a rate of $10/hr.

The Gorilla Experiment Builder (www.gorilla.sc) was used to create and host this and all experiments that follow (Anwyl-Irvine et al., 2020). After consenting to participate in the study, a visual search bot check, in which the participant had to choose the picture of the cat among five pictures of dogs presented in a 3 × 3 grid, was administered. All participants in all experiments passed this check. Each participant was then asked 64 of the general information questions from the pool used in the previous two experiments in a random order (see the online supplemental material). With each question on screen, participants were asked to provide their answer in a box, then on the next screen, they rated their confidence in the correctness of their answer on a sliding scale from 0 (*not at all*) to 100 (*very*). Instead of being asked whether they were curious to know the answer (as in the previous experiments), in the current experiment participants were then asked how surprised they would be if they found that their answer was wrong. They made their responses on a sliding scale from 0 (*not at all*) to 100 (*very*). Presumably, if they strongly believed they were right, they would be very surprised to find out they were wrong. If the second-order judgment was sufficient that people noticed when they were wrong, then they would not be surprised to be wrong (when, indeed they were wrong). Thus, we inverted this score, coding their expectation of (having made an) error as 100 minus their surprise score, and labeled this dependent variable "subjective uncertainty." All responses were self-paced, and participants had to respond to each part of the trial before moving on to the next trial. Once all questions were answered and ratings were made, we provided participants with the correct answers to all the questions.

Results

We hand scored errors, counting as correct items that contained spelling or grammar errors. Seven trials were excluded due to a missing accuracy value. People were correct on 40% (SD = 15%) of the questions. Confidence ratings were strongly correlated with correctness (γM = .69, SD = .15) indicating excellent metacognition (see Wixted & Wells, 2017). This is a result that is typically found in the literature, and provides assurance that participants took the task seriously, despite its being online.

Our measure of second-order metacognition or "expected subjective uncertainty" was operationalized as the complement of their second rating (i.e., 100 minus their expected surprise at being wrong). We analyzed these data as before. Model comparison using WAIC again strongly favored including the quadratic term of confidence.

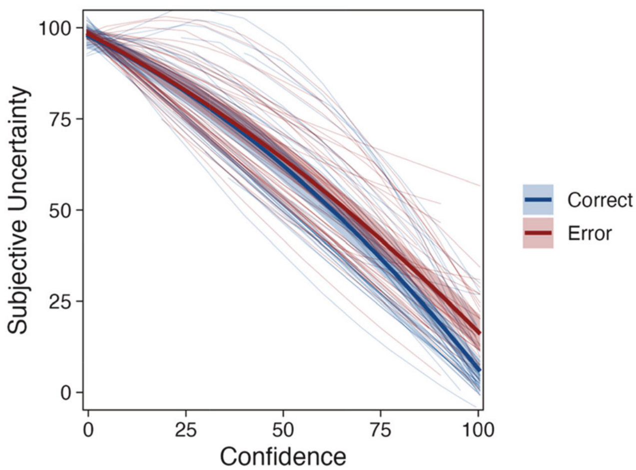
Participants' initial confidence negatively predicted subjective uncertainty for both correct ($b_{\text{Linear}} = -.91$, SE = .02, 95% CI [−.96, −.87], $p_d > 99.9\%$; $b_{\text{Quadratic}} = -.11$, SE = .03, 95% CI [−.16, −.06], $p_d > 99.9\%$) and incorrect answers ($b_{\text{Linear}} = -.82$, SE = .03, 95% CI [−.87, −.76], $p_d > 99.9\%$; $b_{\text{Quadratic}} = -.07$, SE = .02, 95% CI [−.11, −.03], $p_d = 99.9\%$; see Figure 5). The difference between these correlations was significant for the linear comparison and not quite significant with the quadratic (Linear = .10, 95% CI [.04, .15], $p_d = 99.9\%$; Quadratic = .04, SE = .02, 95% CI [.00, .08], $p_d = 96.0\%$). A post hoc comparison on subjective uncertainty at confidence = 90, between correct and incorrect answers indicated that high-confidence errors elicited slightly greater subjective uncertainty than did high-confidence correct answers (Difference = 8.02, SE = 1.84, 95% CI [4.46, 11.70], $p_d > 99.9\%$). As can be seen from Figure 5, even when they made errors, participants, for the most part, stuck to their original level of certainty that they had been correct. There was a small difference between highly confident correct answers and highly confident errors, however, and it was in the same direction as had been observed with the curiosity ratings.

Discussion

The results of Experiment 3 indicated that people expected that they would be slightly less surprised should they be wrong when

Figure 5

Experiment 3: The Relation Between Subjective Uncertainty Ratings and Participants' Original Confidence in Their Responses With No Feedback



Note. See the online article for the color version of this figure.

they had made an error with high confidence, than when they had given a correct response with equally high confidence. These results provide some support for the notion that curiosity ratings– which show a similar, but possibly larger, difference between correct and incorrect highly confident responses– may tap into discrepant information which is ignored in initial confidence judgments (perhaps because of a confirmation bias), but which may be accessible with second-order metacognitive judgments, especially those that specifically target the possibility of being wrong (and see Brewer et al., 2005). Even so, the effects with the subjective uncertainty judgments were small, and could not be directly compared to those observed in curiosity ratings because of differences in procedures and online versus in person participant populations. A direct comparison was implemented in the next experiment.

In addition, several theories (e.g., Friston et al., 2017; Gottlieb et al., 2013; Gottlieb & Oudeyer, 2018; Loewenstein, 1994; Marvin et al., 2020; Murayama et al., 2019) have proposed that curiosity is related to either the reduction of subjective uncertainty or to prediction error. For this reason, a direct contrast of the relation between people's subjective uncertainty and their judgments of curiosity is warranted.

Experiment 4: The Relation Between Curiosity and Expected Subjective Uncertainty

To more directly determine whether curiosity ratings were essentially "nothing more" than subjective uncertainty ratings, as well as to compare the magnitude of the difference between high-confidence errors and high-confidence correct answers for curiosity as contrasted to subjective uncertainty ratings, Experiment 4 was conducted. One group of participants made ratings of the expectation of surprise at the possibility of being wrong (i.e., subjective uncertainty ratings), while the other group made judgments of curiosity.

Method

Ninety-eight adults with an average age of 29 ± 6.7 years completed the study. Forty-five participants identified as female and 51 as male (two did not specify). One participant reported some high school experience, eight participants reported that high school was their highest level of education, 13 reported an associate or professional degree, 22 reported some college, 34 reported having a bachelor's degree, 18 reported having a master's degree, and one reported a doctorate degree (one did not report their level of education).

The procedure was like that of Experiment 3: Participants were asked a series of 64 general information questions (see the online supplemental material), gave their answers, and then rated their confidence in the correctness of each answer on a sliding scale from 0 (*not at all*) to 100 (*very*). No feedback was given. Following the confidence judgment, in a between-participants design which randomly included 42 participants in the subjective uncertainty condition and 56 participants in the curiosity condition, participants were asked after each confidence rating, respectively, either of the following: "How surprised would you be if your answer was wrong?" or "How curious are you to see the correct answer to this question?" Participants made their response to this question on a sliding scale from 0 (*not at all*) to 100 (*very*). As before, we inverted the surprise ratings to ratings of subjective uncertainty. At the end of the experiment participants were allowed to see the correct answers to the general information questions.

Results

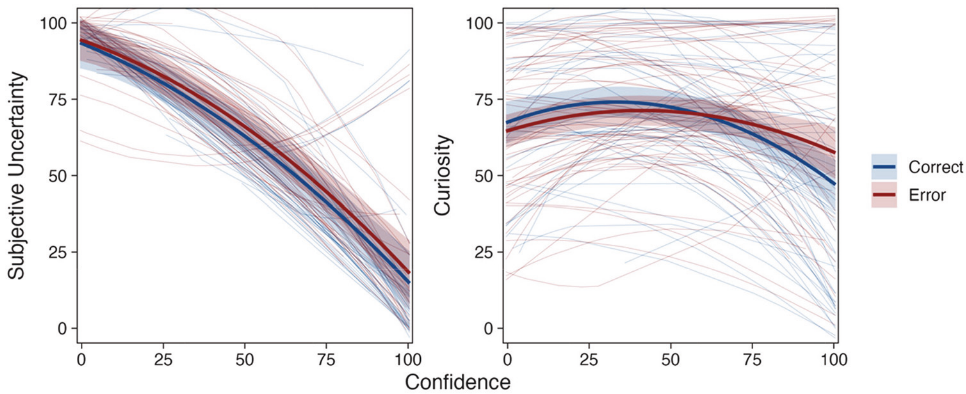
Participants were correct on 47% (SD = 19%) of the questions (two trials were removed due to a missing accuracy value). Confidence ratings were strongly correlated with correctness (γM = .68, SD = .16), indicating excellent metacognition overall.

As can be seen by comparing the left panel of Figure 6 (portraying subjective uncertainty) with the right panel (portraying curiosity), the two kinds of judgments were quite different from one another ($\text{Difference}_{\text{Linear}} = -.63$, SE = .07, 95% CI [−.77, −.50], $p_d > 99.9\%$; $\text{Difference}_{\text{Quadratic}} = .04$, SE = .04, 95% CI [−.04, .11], $p_d = 80.2\%$). Uncertainty ratings showed a sharp and near linear decline with confidence ($b_{\text{Linear}} = -.77$, SE = .05, 95% CI [−.87, −.67], $p_d > 99.9\%$; $b_{\text{Quadratic}} = -.10$, SE = .03, 95% CI [−.15, −.04], $p_d > 99.9\%$), whereas curiosity ratings showed an inverted U-shaped curve overall ($b_{\text{Linear}} = -.14$, SE = .04, 95% CI [−.22, −.05], $p_d > 99.9\%$; $b_{\text{Quadratic}} = -.13$, SE = .03, 95% CI [−.18, −.08], $p_d > 99.9\%$).

We also looked specifically at the difference at high levels of confidence between correct answers and errors for the two kinds of judgments. As the left panel shows, participants' initial confidence negatively predicted subjective uncertainty for both correct ($b_{\text{Linear}} = -.78$, SE = .06, 95% CI [−.89, −.67], $p_d > 99.9\%$; $b_{\text{Quadratic}} = -.09$, SE = .04, 95% CI [−.16, −.02]; $p_{\text{d}} = 99.1\%$) and incorrect answers ($b = -.76$, SE = .06, 95% CI [−.87, −.65], $p_d > 99.9\%$; $b_{\text{quadratic}} = -.10$, SE = .03, 95% CI [−.16, −.04]; $p_{\text{d}} = 99.9\%$). The difference between these correlations was not significant (Linear = .02, SE = .04, 95% CI [−.06, .11], $p_d = 70.5\%$; Quadratic = −.01, SE = .03, 95% CI [−.07, .05], $p_d = 63.3\%$). In contrast, as the right panel shows, when people were asked for their curiosity ratings (on the same items), there was a negative correlation between curiosity

Figure 6

Experiment 4: Subjective Uncertainty Ratings (Left Panel) and Curiosity Ratings (Right Panel) for Correct and Incorrect Items With No Feedback As a Function of Participants' Confidence in Their Original Responses



Note. See the online article for the color version of this figure.

and initial confidence for correct answers ($b_{\text{Linear}} = -.20$, $SE = .05$, 95% CI [−.29, −.11], $p_d > 99.9\%$; $b_{\text{quadratic}} = -.16$, $SE = .03$, 95% CI [−.21, −.10]; $p_d > 99.9\%$), but there was no linear association for errors ($b_{\text{Linear}} = -.07$, 95% CI [−.16, .02], $p_d = 93.7\%$; $b_{\text{Quadratic}} = -.10$, $SE = .03$, 95% CI [−.15, −.05]; $p_d = 99.9\%$). The difference between these two trends was significant (Linear = .13, $SE = .04$, 95% CI [.06, .20], $p_d > 99.9\%$; Quadratic = .05, $SE = .02$, 95% CI [.01, .11], $p_d = 98.5\%$). A post hoc test of differences in curiosity between errors and correct responses, at confidence = 90, showed that high-confidence errors were associated with greater curiosity than were high-confidence correct answers (Difference = 7.22, $SE = 2.33$, 95% CI [3.41, 11.24], $p_d = 99.9\%$). The difference in subjective uncertainty, though, at confidence = 90, was not significant (Difference = 3.36, $SE = 2.33$, 95% CI [−1.15, 7.86], $p_d = 92.2\%$). As Figure 6 also indicates, the three-way interaction between correct/incorrect, linear confidence, and rating (curiosity or subjective uncertainty) was significant ($b = -.03$, $SE = .01$, 95% CI [−.05, .00], $p_d = 97.7\%$).

Discussion

The difference in curiosity observed here between high-confidence errors and high-confidence correct answers—replicating the results of Experiment 2 and the pilot experiment—was significant. People were more curious about their high-confidence errors. The difference between high-confidence errors and high-confidence correct answers with ratings of subjective uncertainty, though, was not significant. These results suggest that curiosity ratings may tap into discrepant information, indicating that an answer might be wrong, to a greater extent than occurs by simply questioning people about the possibility that they could be wrong. Thus, although Experiment 3 had suggested that the difference between correct and incorrect high-confidence responses on the curiosity measure might be attributable to the curiosity query acting, merely, as a second-order metacognitive rating or perhaps to "a Bayesian resampling/regression" (Stephen Spiller, personal communication, April 7, 2021), the present experiment indicates that people's curiosity ratings are more potent in terms of distinguishing between correct responses and errors. Asking if people are curious seems to target some information that distinguishes correct responses from errors, over and above the information that is accessed by a second metacognitive question about subjective uncertainty.

The results of Experiment 4 also provide information about the overall relation between subjective uncertainty and curiosity. Friston et al. (2017) have postulated that curiosity is based on a principle of uncertainty reduction or of free energy minimization: "Resolving uncertainty about the world, through active inference, necessarily entails curious behavior and consequent 'aha' or eureka moments" (Friston et al., 2017, p. 2634). Similarly, a number of theories have proposed a relation between prediction error and curiosity (Gottlieb et al., 2013; Gottlieb & Oudeyer, 2018; Loewenstein, 1994; Murayama et al., 2019). It should be noted that insofar as the assessment of curiosity is prefeedback, there is no way to evaluate the actual error of predictions at the time of curiosity assessment, since the ascertainment of the existence of an error or lack thereof has not yet occurred. The estimates that participants provided, though, provide a proxy as to what their subjective prediction error expectation was. Presumably, these theoretical positions hold that the greater the expected prediction error or subjective uncertainty, the greater should be people's curiosity[2]. Subjective uncertainty, though, as assessed in this experiment, was unrelated to curiosity.

Summary of Part A

Yes/no feedback had a dramatic effect on curiosity for correct answers at all levels of confidence. When people had given the correct answer and were told that they were right, their curiosity

[2] It is also possible that curiosity is associated with the change in subjective uncertainty over time, that is, the rate at which uncertainty is being resolved (see Kornell & Metcalfe, 2006b and Metcalfe et al., 2021 for related views). This can be evaluated, at least over very short intervals, by comparing the original confidence judgment and the second metacognitive judgment concerning subjective uncertainty. However, Experiment 3 and 4 both indicted that there was little change at any level of confidence.

vanished, as expected. When they thought that they were right, that is, they had high confidence, then were told that they were wrong, their curiosity was piqued. If they had had low confidence about their own erroneous responses, being told they were wrong did little to change curiosity, presumably because they already suspected that they were wrong.

Two other findings from Part A are notable. First, curiosity ratings are not tantamount to assessments of subjective uncertainty, free energy, or prediction error: The functions for curiosity and subjective uncertainty were unlike one another. Second, in the absence of feedback, people were more curious to know the answer to high-confidence erroneous responses than to high-confidence correct responses. It is as if the curiosity ratings tapped into some kind of information about the truth of the answer that mere confidence ratings did not access–perhaps a slight feeling of unease or conflict. Even asking people to reassess the possibility that they might be wrong, using second-order metacognitive judgments, did not have as salient a discriminating effect as did curiosity judgments. People's feelings of curiosity, then, distinguished between high-confidence errors and equally high-confidence correct responses.

Part B. The Effect of True But Uncertain Feedback

In Experiments 1 and 2, we tested the effects on curiosity of simply being told, by a reliable source, whether a given answer was right or wrong. When such unambiguous feedback is provided, the effects on curiosity are large and straightforward, as was demonstrated in those experiments. In the real world, though, it is often the case that there is uncertainty associated with the feedback people receive. People obtain information on the Internet, through newspapers, from friends or coworkers, and from a variety of other sources that may or may not be reliable. They often do not know whether that information is actually true and correct (see, e.g., Rich et al., 2017). Their curiosity might be affected, however, even when the truth of the information producing the effect is uncertain. In the following experiments, we investigated the effects of receiving information about the answer under conditions of uncertainty. To do so, we provided participants with the answers purportedly given by another participant. Those answers could be true or false.

In Experiments 5, 6, and 7, we investigated the effects of receiving true but uncertain informational feedback: We provided the correct answer as feedback but under the guise that it was the answer given by another participant and without telling subjects that the answer was, in fact, true. In Experiment 5, we asked participants, after they received this informationally true (but uncertain) feedback about their curiosity to know the answer. In Experiment 6, we investigated their degree of belief in the true answer, so given. In Experiment 7, we investigated how certain they were that they then knew the answer after having been provided with true (but uncertain) feedback. Then we looked at the effects of false but uncertain feedback in a similar manner. In Experiment 8, we investigated curiosity following the receipt of false (but uncertain) feedback. In Experiment 9, we investigated participants' belief in the false feedback. In Experiment 10, we investigated people's certainty that they now knew the answer following receipt of feedback that was false.

To lay the groundwork for comparisons of the effects of true and false (but uncertain) feedback on curiosity about correct and incorrect responses, however, we first conducted a control experiment in which no feedback was given. This allowed us to chart the change in curiosity from no feedback to true feedback under uncertainty, and from no feedback to false feedback under uncertainty. Further, through the use of a slightly different procedure than that used in our previous experiments, it provided an additional possibility for replication of the finding that when no feedback is given people are more curious about high-confidence errors than they are about equally high-confidence correct answers.

Control Experiment: No Feedback Provided

Method

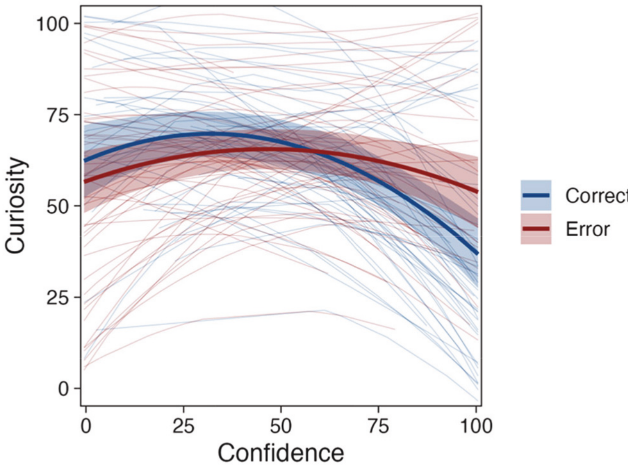
Participants. Fifty participants completed the experiment. The average age was 27.5 ± 5.99 years. Thirty-two participants identified as female, 15 as male, and three as nonbinary. Eight participants reported that high school was their highest level of education, three reported associate's degree, 18 reported college, seven reported master's degree, nine reported some college, and five reported trade school.

Procedure. Participants were instructed that they would be shown a series of general information questions (see the online supplemental material) and that after each question they should type in their answer, guessing if necessary, and that they would then be asked for their confidence that their answer was correct. They were told that at the end of the experiment they would be shown the answers to 10% of the questions, and that they would be given the opportunity to indicate how curious they were to see the answer to each particular question. This instruction follows the procedure of Bloom et al., 2018, which, they argued, resulted in more discriminating curiosity judgments than did merely asking people whether they were curious or not. (It differed from the curiosity conditions of Experiments 1, 2 and 4 but is the same as Experiments 5 and 8 in this regard.) Each participant was asked 64 of the general information questions from the pool used in the previous experiments in a random order. After they answered each question, they rated their confidence in the correctness of their answer on a sliding scale ranging from 0 (*not at all*) to 100 (*very*). They were then immediately asked to indicate their curiosity to see the answer to that particular question, on the same sliding scale, without having received any feedback.

Results

As is shown in Figure 7, the results of the no-feedback experiment indicated that people were more curious about their high-confidence errors than they were about their high-confidence correct responses: Curiosity decreased with increased confidence for correct answers ($b_{\text{Intercept}} = 60.96$, $SE = 3.26$, 95% CI [54.54, 67.15], $p_d > 99.9\%$; $b_{\text{Linear}} = -.25$, $SE = .06$, 95% CI [−.37, −.13], $p_d >$ 99.9%; $b_{\text{Quadratic}} = -.19$, $SE = .04$, 95% CI [−.26, −.10], $p_d >$ 99.9%), but remained relatively stable for errors ($b_{\text{Intercept}} = 59.92$, $SE = 3.09$, 95% CI [53.92, 65.92], $p_d > 99.9\%$; $b_{\text{Linear}} = -.03$, $SE = .06$, 95% CI [−.14, .09], $p_d = 68.2\%$; $b_{\text{Quadratic}} = -.11$, $SE =$ .03, 95% CI [−.16, −.05], $p_d > 99.9\%$). The linear trends were significantly different (.22, $SE = .04$, 95% CI [.14, .31]), and a post hoc test comparing curiosity at confidence = 90 between errors and correct answers indicated that high-confidence errors elicited greater curiosity than did correct answers (Error–Correct = 11.94, $SE = 2.48$, 95% CI [7.01, 16.67], $p_d > 99.9\%$).

Figure 7
Control Experiment: Curiosity As a Function of Confidence for Correct Answers and Errors With No Feedback



Note. See the online article for the color version of this figure.

Discussion

This result replicates the finding that curiosity is greater for high-confidence errors than it is for equally high-confidence correct responses found in Experiment 2, in the pilot to Experiment 2 (see Figure 4, right panel), and in the curiosity condition of Experiment 4. These results also provide the no-feedback baseline for the true or false feedback given in the following Experiments 5 and 8, respectively.

Experiment 5: Effects of True Feedback Under Uncertainty on Curiosity About Correct Answers and Errors

In Experiment 5, we again investigated the effect of feedback on curiosity. Unlike in Experiment 2, however, where definitive yes/no feedback was provided, in Experiment 5 we provided informational feedback and told participants that the provided answer was the answer that had been given by another participant. Thus, rather than being told definitively that their answer was right or wrong, the correct answer to each question was provided under conditions of uncertainty. We investigated how or if this correct information swayed or sated participants' curiosity.

Method

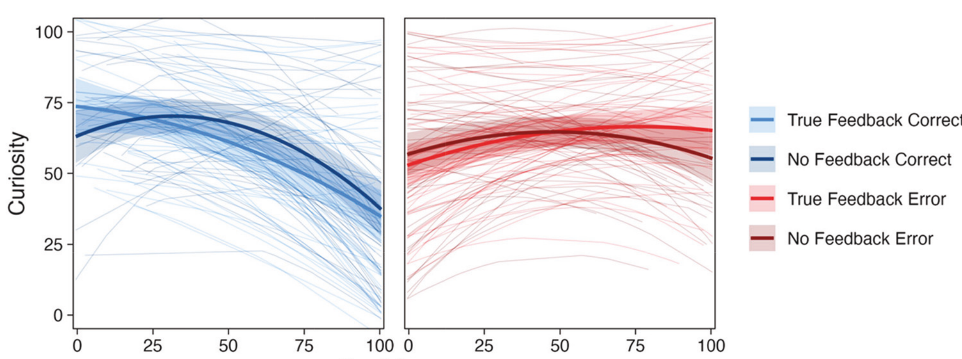
Participants. Fifty participants completed the experiment. The average age was 24.71 ± 4.59 years (one participant did not report). Thirty-eight participants identified as female, nine as male, and three as nonbinary. Eight participants reported that high school was their highest level of education, one reported associate's degree, 18 reported bachelor's degree, seven master's degree, 13 some college, two trade school, and one declined to report.

Procedure. The procedure was identical to that of the control experiment, except that immediately after making their confidence judgment about their own answer, participants were shown an answer (which was factually correct) and were told that another participant had given it as the answer. After being shown this answer, they were then asked to say how curious they were to "see the correct answer at the end." Questions and answers are provided in the online supplemental material.

Results

The results from Experiment 5 were compared with the control (no-feedback) experiment described earlier by fitting a model to the data from both experiments. The control (no-feedback) experiment compared with Experiment 5 was coded as a dummy factor, and interactions with it and all other effects were included. As is shown in Figure 8 (left panel), there was no difference in the linear association, though there was a small difference in the quadratic association, between confidence and curiosity for correct responses when true feedback was provided as compared with when no feedback was given (true feedback vs. no feedback = −.13, *SE* = .08, 95% CI [−.29, .03], p_d = 94.5%; quadratic difference = .11, *SE* = .06, 95% CI [.00, .22], p_d = 98.0%). The post hoc analyses comparing curiosity at three confidence levels were as follows: confidence = 10, true feedback vs. no feedback = 5.41, *SE* = 5.41,

Figure 8
Experiment 5: Curiosity Following the Provision of True, Uncertain Feedback Compared With No Feedback As a Function of Participants' Initial Confidence in Their Correct Answers (Left Panel) or Their Errors (Right Panel)



Note. See the online article for the color version of this figure.

95% CI [−5.01, 16.06], p_d = 84%; confidence = 50: −6.80, *SE* = 4.48, 95% CI [−15.55, 2.07], p_d = 93.7%; confidence = 90: −5.14, *SE* = 5.61, 95% CI [−15.66, 6.38], p_d = 81.7%).

As Figure 8 (right panel) shows, provision of the correct answer under uncertainty, as compared to no feedback, had a slight effect on errors: the linear confidence-curiosity association was greater with true feedback than no feedback (difference in slopes for no feedback versus true feedback: .14, *SE* = .07, 95% CI [.01, .27], p_d = 97.8%; quadratic: .03, *SE* = .04, 95% CI [−.05, .10], p_d = 77.1%;. The post hoc analyses on errors were as follows: at confidence 10 [true feedback vs. no feedback]: −3.40, *SE* = 4.74, 95% CI [−12.30, 6.00], p_d = 76.1%; at confidence 50: .30, *SE* = 3.80, 95% CI [−7.11, 7.59], p_d = 53.1%; at confidence 90: 7.47, *SE* = 5.26, 95% CI [−2.59, 17.87], p_d = 92.3%).

Discussion

It is plausible that people might have believed the other participant's answer to be true and considered the possibility that their own high-confidence answer was wrong when they were given an answer that was different from their own and that was, in fact, correct. Such considerations might have been responsible for slightly increasing their curiosity, in a manner similar to the effect of yes/no feedback in which people became more curious when they were told that the answer they had just given with high confidence was wrong. However, the effects of true but uncertain feedback were small. Further, this very small effect was quite different from the large effect observed when yes/no feedback was given. With the later, people's curiosity decreased sharply, almost to zero when they were right. In contrast, here, there was little to no change in curiosity: merely telling the participant that someone else gave the same answer was insufficient to push participants over the hard boundary illustrated in the left panel of Figure 1: they still did not know for sure whether their answer was correct or not (perhaps the other participant was wrong), and hence their curiosity was not extinguished.

The preceding explanation of the small effects of uncertain but true feedback is predicated on assumptions about participants' belief in whether the answers that were given by the other participant were true or not. Accordingly, in the next experiment, we directly probed participants' belief in the feedback.

Experiment 6: Belief in True Feedback Under Uncertainty

Experiment 6 was identical to Experiment 5, except that instead of asking participants for their curiosity ratings after they had received an answer purportedly given by another participant, they were asked to indicate how sure they were that the other participant's answer was correct.

Method

Forty-eight participants completed the study. They averaged 23.00 ∓ 7.50 years (two participants did not report age). Twenty-nine participants identified as female, and 19 as male. Twelve participants reported that high school was their highest level of education, one reported bachelor's degree, 34 some college, and one trade school. In order to offset the possibility that participants would infer that all of the feedback was correct, we included 6 items that were accompanied by false feedback: two false feedback items were given at the beginning of the experiment; two were given one third of the way through; two were given two thirds of the way through (see the online supplemental material for all items used). We only analyzed the results for items that were accompanied by true feedback.

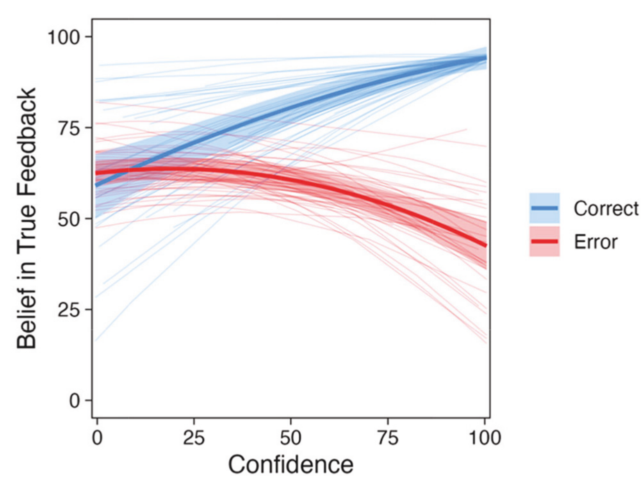
Results

Participants were correct on 32% (*SD* = 17%) of the questions (we dropped one trial due to a missing accuracy value); confidence ratings were strongly correlated with correctness (γM = .74, *SD* = .14) and participants were overconfident; their original confidence judgments averaged 43% (*SD* = 17%).

As is shown in Figure 9, it mattered for their belief in the feedback whether participants, themselves, had given the correct answer or had made an error (b_{Accuracy} = −10.66, *SE* = .98, 95% CI [−12.58, −8.71], $p_d > 99.9\%$). When participants had been correct, their belief in the (true) answer given by the other participant was very high when they had originally been highly confident about their own answer. For items other than their extremely high confident responses, however, belief in the answer given by another participant was not close to ceiling, indicating that participants had less than perfect confidence in this answer (b_{Linear} = .35, *SE* = .04, 95% CI [.27, .44], $p_d > 99.9\%$), just as they had had less than perfect confidence about their own answers.

When participants had made an error, they exhibited a moderate amount of belief that the answer given by the other participant (which was different from their own response) was correct, regardless of whether they had had high or low confidence in their own answer. Notably, even when participants expressed high confidence about their own answer, they acknowledged considerable belief that the answer given by the other participant was correct (b_{Linear} = −.20, *SE* = .03, 95% CI [−.26, −.13], $p_d > 99.9\%$).

Figure 9
Belief in the True Answer Given as Feedback Under Uncertainty, Depending on Correctness or Incorrectness of Original Answer, and As a Function of Their Initial Confidence in Their Answer



Note. See the online article for the color version of this figure.

Discussion

The data from Experiment 6 are consistent with the idea that true feedback under uncertainty does little to quench people's curiosity when they had been correct, presumably because participants remained uncertain about any but their highest confidence responses. These data are also consistent with the idea that the slight increase in curiosity for high-confidence errors might have resulted from doubt flagging instilled by the feedback that they thought might be correct (which implied that they were wrong). Rather than showing disbelief in the conflicting feedback, they showed considerable belief, even when they had initially had high confidence that their own wrong answers were correct. If they had rejected these answers when they thought their own answers were right the pattern for errors would have tended to zero as people's confidence in their own answers increased. Instead, they showed considerable belief in the correct answer when they had made high-confidence errors. Although these data are generally consistent with the RPL model, which proposes that curiosity is based on people's evaluation of whether they know what the answer is, or instead only "almost known." Thus, in Experiment 7, following true but uncertain feedback, we asked them directly whether they now knew the answer after they had seen the feedback.

Experiment 7: Feeling That They Knew the Answer After Receiving True Feedback Under Uncertainty

In Experiment 7, we investigated the change in people's confidence in knowing the answer after being exposed to true (but uncertain) feedback. We expected that participants would still not be completely sure that they knew the answer when they had been correct, for any but their highest confidence responses. For high-confidence errors, we expected that the correct answer feedback under uncertainty might undermine their confidence that they now knew the answer: instead of being sure they knew, they might now think that they only "almost knew" for some high-confidence errors. The procedure was like that of Experiment 5, except instead of being asked for their curiosity to know the answer, participants were asked, immediately following the feedback, how certain they were that they now knew the correct answer.

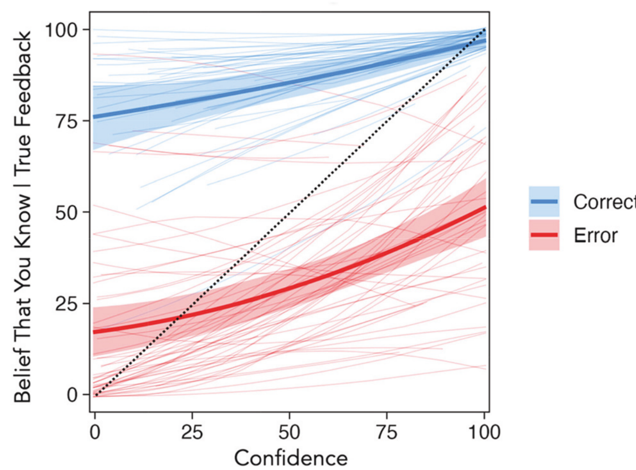
Method

Fifty-four participants completed the study. Participants averaged 27.21 ∓ 5.56 years (two participants did not report age). Twenty-nine participants identified as female, 23 as male, and two as nonbinary. Three participants reported that high school was their highest level of education, 23 reported bachelor's degree, 13 reported some college, four associate's degree, eight master's degree, and two trade school.

Results

As is shown in Figure 10, when participants' had been correct in their answer, their certainty about knowing the true answer increased, but did not reach ceiling, for low confidence corrects, and it remained very high for high-confidence corrects ($b_{\text{Intercept}}$ = 84.71, SE = 2.28, 95% CI [80.03, 89.02], $p_d > 99.9\%$; b_{Linear} = .21, SE = .04, 95% CI [.12, .29], $p_d > 99.9\%$). When they had given the wrong answer, their confidence that they now knew the correct answer decreased for high-confidence errors and remained low for low confidence errors ($b_{\text{Intercept}}$ = 29.90, SE = 2.68, 95% CI [24.82, 35.40], $p_d > 99.9\%$; b_{Linear} = .34, SE = .05, 95% CI [.25, .43], $p_d > 99.9\%$).

Figure 10
Experiment 7: Participants' Assessment of Knowing the Answer After Receiving True (But Uncertain) Feedback As a Function of Their Initial Confidence, Depending on the Correctness or Incorrectness of Their Answer



Note. The dotted black diagonal line indicates no change. See the online article for the color version of this figure.

Discussion

Only when participants had initially expressed the very highest level of confidence in the truth of their own answer and were also given true feedback, were they convinced that they knew the correct answer. The finding that, even when given true feedback, participants often were unconvinced that they now knew the true answer even when they, themselves, had been correct initially, seems sufficient to account for their curiosity not being sated by true-answer feedback. It would seem that knowledge of truth has to be definitive to have a sating effect. Similarly, the effect of true-answer feedback when people had made high-confidence errors was a decrease in belief that they now knew the answer. That decrease might have resulted in their increased curiosity for those answers, as was shown in Experiment 5. Under such circumstances, the discrepant feedback may have flagged some doubt and stimulated curiosity, in a manner similar to what occurred upon finding out that they had been wrong when they thought they had been right in Experiment 2.

Summary of Part B

When people had been correct and were told that another participant produced the same answer that they had, their curiosity changed very little. Although participants belief in the correct answer increased for all but the highest confidence correct answers, it did not reach full certainty. Participants' evaluation that they knew the answer after being provided with feedback, also, did not reach the highest levels of certainty for any but their highest confidence correct answers. Thus, because they did not fully believe that they knew the answer after getting feedback, it is likely that true but uncertain feedback did not have the same extinguishing effect on curiosity that being simply told, explicitly, that they were right had in Experiment 2.

True but uncertain feedback did change curiosity about high-confidence errors, though the effect was small. The discrepant (but true) answer given by the other participant decreased participants' confidence that they now knew the answer and increased their curiosity. Participants presumably recognized that there was a chance that they had been wrong initially. The true but uncertain feedback had a similar effect on curiosity as had the yes/no feedback: upon realizing that they were, or might be wrong, curiosity about high-confidence erroneous responses increased.

Finally, it is worth noting that the feedback given in this set of experiments was purportedly provided by a nonauthoritative source (another participant), and repeated only once. Our results might have been quite different if the source had been the Trivial Pursuits World Champion, or the Encyclopedia Britannica—a fully trustworthy source (see Lewandowsky et al., 2012). The results might also have been different if the source had been a beloved in-group or an antagonistic out-group member, for instance. Further research is needed to understand how the characteristics of the conveyor of information affects curiosity. However, as demonstrated here, the mere presentation of true but uncredentialled information had surprisingly small effects on curiosity.

Part C. The Effect of False But Uncertain Feedback

The final set of experiments was nearly identical to those in Part B, except that false feedback, rather than true feedback, was given.

Experiment 8: Effect of False Feedback Under Uncertainty on Curiosity for Correct Answers and Errors

Method

This experiment was identical to Experiment 5 except that the feedback (the response given by the other participant) was a false, but plausible, answer (e.g., "The capital of Canada is Toronto") taken from the materials of Metcalfe and Eich (2019; see also the online supplemental material). Forty-six participants took part and averaged 27.87 ± 6.66 years (two respondents did not report age). Twenty-two participants identified as female, 20 as male, one as gender fluid, two as nonbinary, and one declined. Six participants reported that high school was their highest level of education, 11 reported bachelor's degree, 12 some college, two associate's degree, nine master's degree, three trade school, one eighth grade, one none, and one professional degree.

Results

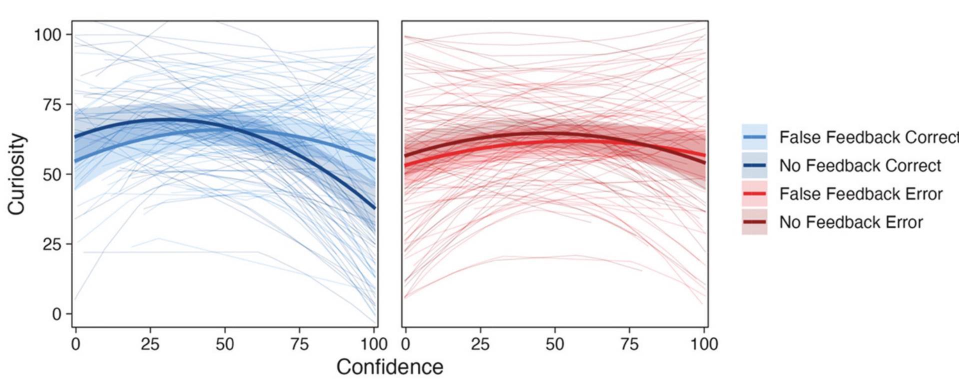
As can be seen from the left panel of Figure 11, false feedback under uncertainty had an effect on people's curiosity when they had been correct. The linear association between confidence and curiosity for correct answers was greater in the feedback than in the no-feedback condition (difference in slopes [false feedback vs. no feedback]: .26, SE = .08, 95% CI [.09, .41], p_d = 99.9%). False feedback especially increased people's curiosity about their high-confidence correct responses (at confidence 10 [false feedback vs. no feedback]: −7.93, SE = 5.44, 95% CI [−18.75, 2.57], p_d = 92.6%; at confidence 50: −1.14, SE = 4.16, 95% CI [−8.85, 7.69], p_d = 61.8%; at confidence 90: 12.46, SE = 5.17, 95% CI [2.67, 22.76], p_d = 99.2%).

False feedback had little, overall, effect on curiosity about incorrect answers. There was no difference in the linear association between confidence and curiosity for incorrect items between the experiments (difference in slopes [false feedback vs. no feedback]: .06, SE = .07, 95% CI [−.07, .20], p_d = 80.5%). Post hoc tests comparing curiosity for incorrect items between the no feedback and the false feedback conditions were not significant at confidence = 10 [false feedback vs. no feedback]: −3.66, SE = 4.50, 95% CI [−12.30, 5.46], p_d = 79.7, at confidence = 50: −2.85, SE = 3.72, 95% CI [−10.11, 4.77], p_d = 78.5%, or at confidence = 90: 1.18, SE = 5.50, 95% CI [−9.06, 12.41], p_d = 57.8%).

In this experiment, in which the false feedback that was provided was plausible, there was a possibility that when people had committed an error, the false feedback provided could have matched the error that the person had just made (e.g., the participant erroneously said that the capital of Canada is Toronto and

Figure 11

Experiment 8: The Effect of False Feedback on Curiosity



Note. Shown is curiosity following the provision of false but uncertain feedback as compared with no feedback, depending on participants' initial confidence in their correct answers (left panel) or their errors (right panel). See the online article for the color version of this figure.

was provided with "Toronto" as feedback). To evaluate the possibility that this had a different effect on curiosity than did false feedback that did not match the participant's own error, the error data were separated into cases in which the erroneous responses given by participants were the same and different from the false feedback. Curiosity was not significantly different, overall, between these two error-to-feedback match or mismatch types (−1.12, SE = 1.26, 95% CI [−3.48, 1.44], p_d = 81.8%), but the linear *increase* in the confidence–curiosity association was significantly greater when the participant's response and false feedback did not match (.11, SE = .04, 95% CI [.04, .18], p_d = 99.8%). This difference suggests that it is the discrepancy between the feedback and the high-confidence response that provokes curiosity. We will leave replication of this interesting finding to future research.

Discussion

False feedback had an effect on people's curiosity about their own high-confidence correct responses. When shown the discrepant response given by the other participant, even though they had been correct, people became curious to know the answer. Provision of false but uncertain information had little overall effect on people's curiosity for errors. However, when the erroneous feedback matched their own error they had less curiosity to find out the truth than they had when the erroneous feedback failed to match their own high-confidence error. In the two cases where the false feedback highlighted a discrepancy—when it mismatched with their own correct response and also when it mismatched with their own error—people became more curious to know what the correct answer was. Presumably, the change in curiosity about people's high-confidence correct responses occurred because the false feedback undermined people's feeling that their own correct (or, indeed incorrect) answer was correct.

Experiment 9: Belief in False Feedback Under Uncertainty

As before, people's belief in the false feedback provided by another participant was potentially important in determining whether its provision increased curiosity or not. This belief in the false feedback (see, Marsh & Yang, 2018) was evaluated in Experiment 9.

Method

Forty-eight participants, average age 25.34 ± 6.77 years (two participants did not report age), completed Experiment 9. Twenty-two participants identified as female, 23 as male, two as nonbinary, and 1 declined to respond. Twenty participants reported that high school was their highest level of education, 26 some college, one associate's degree, and one declined to say. One trial was excluded due to a missing accuracy value. To offset the possibility that participants would infer that all the feedback was incorrect, half of the items were accompanied by false feedback and half by true feedback (see the online supplemental material). We only analyzed the responses accompanied by the false feedback.

Results

The results of this experiment indicated that people expressed less belief in the false feedback when they had been correct, as compared to when they had made an error (b_{Accuracy} = 7.79, SE = 1.16, 95% CI [5.57, 10.09], $p_d > 99.9\%$; see Figure 12). The confidence (linear) by accuracy interaction was significant (.19, SE = .04, 95% CI [.12, .27], $p_d > 99.9\%$), such that belief decreased as a function of confidence for correct answers but not for errors.

Discussion

When participants had been wrong, initially, they appeared to simply not know whether to believe the false information that was provided under uncertainty. Since they were agnostic about the information, it had little impact on their curiosity. Similarly, when they had been uncertain about their own correct answer, they appeared to not know whether the false information was true. Interestingly, although they tended to discredit the false feedback with increasing confidence in their own correct answers, they did not entirely reject that feedback as possibly being true, even when they had been confident about their own correct responses. Presumably, this small amount of belief in the false information was sufficient for them to question their own correct response and increase their curiosity about their own high-confidence correct responses.

Experiment 10: Feeling That They Knew the Answer After Receiving False Feedback Under Uncertainty

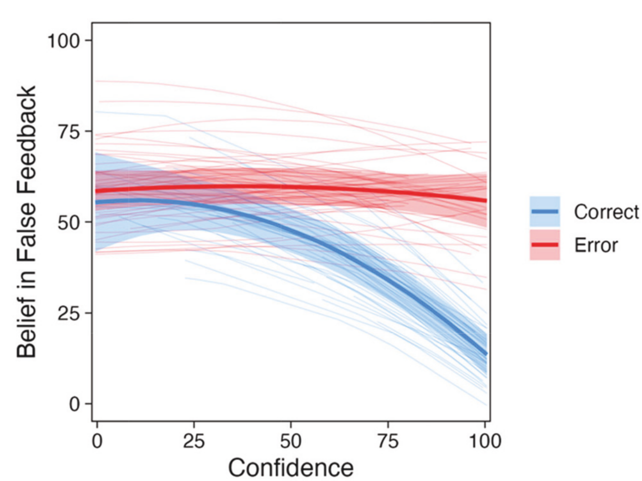
This final experiment investigated participant's evaluation of whether they now knew the correct answer after having received false feedback under uncertainty. Of particular interest was the possibility that participants' feelings that they now knew the answer after receiving false feedback would decrease for their own high-confidence correct responses.

Method

Forty-six participants completed the final experiment. Participants averaged 27.87 ± 6.66 years (two participants did not report age). Twenty-two participants identified as female, 20 as

Figure 12

Experiment 9: Belief in the False Feedback Given Under Uncertainty As a Function of Participants' Initial Confidence in Their Correct or Incorrect Answers



Note. See the online article for the color version of this figure.

male, one as gender fluid, two as nonbinary, and one declined to report. Six participants reported that high school was their highest level of education, 11 reported bachelor's degree, 12 reported some college, two reported associate's degree, nine master's degree, three trade school, one reported eighth grade, one reported "no schooling," and one reported having a professional degree. A total of 64 questions were used in this experiment (see the online supplemental material).

Results

The results indicated that people's belief that they now knew the correct answer following false feedback was altered very little from what it had been before receiving the feedback (b_{Accuracy} = 1.60, SE = .83, 95% CI [−.04, 3.20], p_d = 97.3%; see Figure 13). Their initial confidence in their answer was positively associated with their confidence in knowing the answer following false feedback, for both errors and correct answers (b_{Linear} = .64, SE = .07, 95% CI [.50, .77], $p_d > 99.9\%$).

Discussion

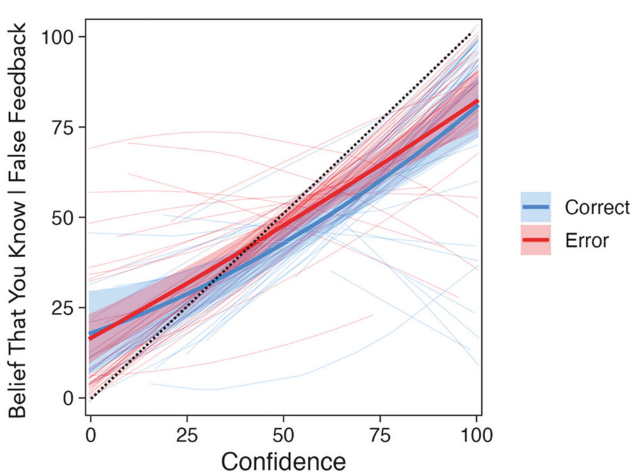
People's confidence in knowing the answer was impacted very little by the receipt of false feedback. There was, however, a slight decline in confidence for responses–both correct and incorrect– that had initially been held with high confidence. The effect of the provision false feedback was dramatically different from that seen in Experiment 7 in which true feedback was given (as shown in Figure 9). Although true feedback following correct answers considerably increased people's belief that they now knew the answers in Experiment 7, false feedback did little to their belief that they now knew.

Summary of Part C

False feedback made people more curious when they had been correct with high confidence. It did not increase curiosity when it matched their own high-confidence errors, though it did when it mismatched their own high-confidence errors.

Figure 13

Experiment 10: Participants' Assessment of Knowing the Answer After Receiving False (But Uncertain) Feedback As a Function of Their Initial Confidence and Depending on the Correctness or Incorrectness of Their Answer



Note. The dotted black diagonal line indicates no change. See the online article for the color version of this figure.

Conclusion

These results, overall, support the idea that curiosity is associated with almost, but not quite, knowing an answer to a question. When people gave a correct answer and were provided with definitive information that they were, indeed, correct, their curiosity dropped to near-zero. Furthermore, feedback manipulations (e.g., factual truth) had little effect on low confidence responses. In addition, when people were correct and they were highly confident that they were correct, they were not very curious. Presumably, under these conditions, they knew that they knew.

People showed increases in curiosity only when they expressed relatively high confidence that they were correct but feedback manipulations undermined that confidence. This, we argue, put them back into their RPL as when (a) they had high confidence that they had been right, but they were told that they were wrong; (b) they had high confidence that they were right (when they were wrong) and got feedback that another participant had a different answer (which was correct); (c) they had high confidence that they were right (when they were wrong) and got feedback that another participant had a different answer (which was wrong); and (d) they had high confidence that they were right (when they were right) but got feedback that another participant had a different answer (which was wrong). In short, when their certainty that their answer was correct was undermined, they became more curious.

These results are broadly consistent with a number of views of curiosity (e.g., Berlyne, 1954; Gruber, Gelman, & Ranganath, 2014; Lau et al., 2020; Loewenstein, 1994; McGillivray, Murayama & Castel, 2015; Murayama, 2021) as well as with the RPL model of curiosity (Kornell & Metcalfe, 2006b; Metcalfe & Jacobs, in press; Metcalfe et al., 2020). People most want to know what the true answer is when they think, metacognitively, that they almost know. When they think they are far from knowing, they are less curious. When they know for sure that they know, they are not curious at all.

A surprising result that emerged from this series of experiments was that curiosity for high-confidence errors was higher than it was for equally high-confidence correct responses. This finding was replicated four times in the present series of experiments. This result did not appear to be solely the result of curiosity judgments being second-order metacognitive judgments in which subjective uncertainty was appraised more thoroughly. Second-order subjective uncertainty judgments showed only a hint of the difference between correct and incorrect responses shown by curiosity ratings. The explanation for why curiosity ratings tap into this distinction between correct answers and errors awaits further research and scrutiny.

Even so, it is possible to speculate that curiosity ratings may be particularly sensitive to discrepancies that indicate that something is amiss. People are curious about slightly unsettling riddles and stumpers (Bar-Hillel et al., 2019), insight problems (Gick & Lockhart, 1995; Kounios & Beeman, 2014), fantastic facts (Fastrich et al., 2018), and magic tricks (Danek et al., 2014; Lau et al., 2020). People might fluently perceive the woman being sawn in half, the kerchief turning into a rabbit, and a ball being pulled from a man's ear. But they know that something beyond that fluent perception must be occurring, indicating that the obvious might not be true. Berlyne (1954) wrote extensively about the notion that curiosity is associated with some kind of "conflict." Furthermore, as detailed by Gruber et al. (2014) in their neural model of curiosity, a primary neural correlate of conflict–the anterior cingulate cortex–is implicated in nearly all studies of curiosity (e.g., Jepma et al., 2012; Ligneul et al., 2018; Morales et al., 2018; Oosterwijk et al., 2020). The feeling of curiosity, from this perspective, may not be a pure measure of only metacognitively perceived closeness to the goal of knowing. The feeling of conflict, or the intuition that something may be amiss, may be an additional component that contributes to curiosity. Indeed, in the experiments presented here, only conflicting feedback (yes/no information that a participant did not know the answer when they thought they did, or information that someone else disagreed with their answer) increased curiosity. Investigations of confirmation biases in confidence judgments (see, Nickerson, 1998; Schwarz, 2015) suggest that people neglect conflicting information when making those judgments. Evaluation of conflict, however, may be intrinsic to assessments of curiosity. The fact that one is curious, then, even though one's confidence in being correct is very high, may be a "tell." That "tell" might provide a clue that allows the discrimination of high-confidence errors from equally high-confidence correct responses.

Context Statement

The RPL model posits that people are most curious when they "almost know." But how do they know that they "almost know"? Presumably they use a retrieval cue (i.e., the question they are asked to answer) to retrieve information. By one widely accepted view, people are highly confident if a great deal of information comes to mind fluently. If little information is retrieved or if it comes to mind slowly, people are less confident. By this view, "almost knowing" entails incomplete information coming to mind quickly. However, "almost knowing" could come about not only because of a slight lack of information or from slightly slow retrieval, but also via the recognition of the presence of conflicting information. Our data suggest that curiosity assessments may be more sensitive to this discrepant information than are confidence ratings. This proposed difference in the weighting of the kind of information invoking curiosity as compared with that on which confidence judgments may be primarily based deserves further empirical investigation. The consequence of curiosity being stimulated by the feeling of conflict-based "almost knowing" would be further information-seeking, perhaps with an emphasis on finding potential flaws or errors. There are likely to be practical implications. If conflicting information triggers curiosity, then discrepant information is, in the context of social media, click bait: It can be used to stimulate curiosity. Such triggering could be used to good end, of course, as when true, but discrepant, information can be used to correct errors, ensure quality, and foster creativity. However, as was shown here, false information can also incite curiosity, even when people were correct. The consequences of this may sometimes be pernicious. The dynamics of curiosity—the conditions under which curiosity occurs, its consequences, and how it is stimulated—deserve further scientific investigation.

References

Anwyl-Irvine, A. L., Massonnié, J., Flitton, A., Kirkham, N., & Evershed, J. K. (2020). Gorilla in our midst: An online behavioral experiment builder. *Behavior Research Methods*, *52*(1), 388–407. https://doi.org/10.3758/s13428-019-01237-x

Asp, E., Manzel, K., Koestner, B., Denburg, N. L., & Tranel, D. (2013). Benefit of the doubt: A new view of the role of the prefrontal cortex in executive functioning and decision making. *Frontiers in Neuroscience*, *7*, Article 86. https://doi.org/10.3389/fnins.2013.00086

Atkinson, R. C. (1972). Optimizing the learning of a second-language vocabulary. *Journal of Experimental Psychology*, *96*(1), 124–129. https://doi.org/10.1037/h0033475

Bar-Hillel, M., Noah, T., & Frederick, S. (2019). Learning psychology from riddles: The case of stumpers. *Judgment and Decision Making*, *13*(1), 112–122.

Berlyne, D. E. (1954). A theory of human curiosity. *British Journal of Psychology*, *45*(3), 180–191. https://doi.org/10.1111/j.2044-8295.1954.tb01243.x

Bloom, P. A., Friedman, D., Xu, J., Vuorre, M., & Metcalfe, J. (2018). Tip-of-the-tongue states predict enhanced feedback processing and subsequent memory. *Consciousness and Cognition*, *63*, 206–217. https://doi.org/10.1016/j.concog.2018.05.010

Brewer, W. F., Sampaio, C., & Barlow, M. R. (2005). Confidence and accuracy in the recall of deceptive and nondeceptive sentences. *Journal of Memory and Language*, *52*(4), 618–627. https://doi.org/10.1016/j.jml.2005.01.017

Brooks, G., Yang, H., & Köhler, S. (2021). Feeling-of-knowing experiences breed curiosity. *Memory*, *29*(2), 153–167. https://doi.org/10.1080/09658211.2020.1867746

Brown, A. S. (1991). A review of the tip-of-the-tongue experience. *Psychological Bulletin*, *109*(2), 204–223. https://doi.org/10.1037/0033-2909.109.2.204

Buratti, S., Allwood, C. M., & Kleitman, S. (2013). First- and second-order metacognitive judgments of semantic memory reports: The influence of personality traits and cognitive styles. *Metacognition and Learning*, *8*(1), 79–102. https://doi.org/10.1007/s11409-013-9096-5

Bürkner, P.-C. (2017). brms: An R package for Bayesian multilevel models using stan. *Journal of Statistical Software*, *80*(1), 1–28. https://doi.org/10.18637/jss.v080.i01

Butterfield, B., & Mangels, J. A. (2003). Neural correlates of error detection and correction in a semantic retrieval task. *Cognitive Brain Research*, *17*(3), 793–817. https://doi.org/10.1016/S0926-6410(03)00203-9

Butterfield, B., & Metcalfe, J. (2001). Errors committed with high confidence are hypercorrected. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*(6), 1491–1494. https://doi.org/10.1037/0278-7393.27.6.1491

Butterfield, B., & Metcalfe, J. (2006). The correction of errors committed with high confidence. *Metacognition and Learning*, *1*(1), 1556–1623. https://doi.org/10.1007/s11409-006-6894-z

Cyr, A. A., & Anderson, N. D. (2013). Updating misconceptions: Effects of age and confidence. *Psychonomic Bulletin & Review*, *20*(3), 574–580. https://doi.org/10.3758/s13423-012-0357-0

Danek, A. H., Fraps, T., von Müller, A., Grothe, B., & Öllinger, M. (2014). Working wonders? Investigating insight with magic tricks. *Cognition*, *130*(2), 174–185. https://doi.org/10.1016/j.cognition.2013.11.003

de Leeuw, J. R. (2015). jsPsych: A JavaScript library for creating behavioral experiments in a web browser. *Behavior Research Methods*, *47*(1), 1–12. https://doi.org/10.3758/s13428-014-0458-y

Dunlosky, J., Serra, M. J., Matvey, G., & Rawson, K. A. (2005). Second-order judgments about judgments of learning. *The Journal of General Psychology*, *132*(4), 335–346. https://doi.org/10.3200/GENP.132.4.335-346

Eich, T. S., Stern, Y., & Metcalfe, J. (2013). The hypercorrection effect in younger and older adults. *Neuropsychology, Development, and Cognition. Section B, Aging, Neuropsychology and Cognition*, *20*(5), 511–521. https://doi.org/10.1080/13825585.2012.754399

Fastrich, G. M., Kerr, T., Castel, A. D., & Murayama, K. (2018). The role of interest in memory for trivia questions: An investigation with a large-scale database. *Motivation Science*, *4*(3), 227–250. https://doi.org/10.1037/mot0000087

Fazio, L. K., & Marsh, E. J. (2010). Correcting false memories. *Psychological Science*, *21*(6), 801–803. https://doi.org/10.1177/0956797610371341

FitzGibbon, L., Lau, J. K. L., & Murayama, K. (2020). The seductive lure of curiosity: Information as a motivationally salient reward. *Current Opinion in Behavioral Sciences*, *35*, 21–27. https://doi.org/10.1016/j.cobeha.2020.05.014

Friston, K. J., Lin, M., Frith, C. D., Pezzulo, G., Hobson, J. A., & Ondobaka, S. (2017). Active inference, curiosity, and insight. *Neural Computation*, *29*(10), 2633–2683. https://doi.org/10.1162/neco_a_00999

Gick, M. L., & Lockhart, R. S. (1995). Cognitive and affective component of insight. In R. J. Sternberg & J. E. Davidson (Eds.), *The nature of insight* (pp. 197–228). MIT Press.

Gilbert, D. T. (1991). How mental systems believe. *American Psychologist*, *46*(2), 107–119. https://doi.org/10.1037/0003-066X.46.2.107

Gilbert, D. T., Krull, D. S., & Mulone, P. S. (1990). Unbelieving the unbelievable: Some problems in the rejection of false information. *Journal of Personality and Social Psychology*, *59*(4), 601–613. https://doi.org/10.1037/0022-3514.59.4.601

Gottlieb, J., & Oudeyer, P. Y. (2018). Towards a neuroscience of active sampling and curiosity. *Nature Reviews Neuroscience*, *19*(12), 758–770. https://doi.org/10.1038/s41583-018-0078-0

Gottlieb, J., Oudeyer, P. Y., Lopes, M., & Baranes, A. (2013). Information-seeking, curiosity, and attention: Computational and neural mechanisms. *Trends in Cognitive Sciences*, *17*(11), 585–593. https://doi.org/10.1016/j.tics.2013.09.001

Gruber, M. J., Gelman, B. D., & Ranganath, C. (2014). States of curiosity modulate hippocampus-dependent learning via the dopaminergic circuit. *Neuron*, *84*(2), 486–496. https://doi.org/10.1016/j.neuron.2014.08.060

Hays, M. J., Kornell, N., & Bjork, R. A. (2013). When and why a failed test potentiates the effectiveness of subsequent study. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*(1), 290–296. https://doi.org/10.1037/a0028468

Jepma, M., Verdonschot, R. G., van Steenbergen, H., Rombouts, S. A., & Nieuwenhuis, S. (2012). Neural mechanisms underlying the induction and relief of perceptual curiosity. *Frontiers in Behavioral Neuroscience*, *6*, Article 5. https://doi.org/10.3389/fnbeh.2012.00005

Kang, M. J., Hsu, M., Krajbich, I. M., Loewenstein, G., McClure, S. M., Wang, J. T., & Camerer, C. F. (2009). The wick in the candle of learning: Epistemic curiosity activates reward circuitry and enhances memory. *Psychological Science*, *20*(8), 963–973. https://doi.org/10.1111/j.1467-9280.2009.02402.x

Kelley, C. M., & Jacoby, L. L. (1996). Adult egocentrism: Subjective experience versus analytic bases for judgment. *Journal of Memory and Language*, *35*(2), 157–175. https://doi.org/10.1006/jmla.1996.0009

Kelley, C. M., & Lindsay, D. S. (1993). Remembering mistaken for knowing: Ease of retrieval as a basis for confidence in answers to general knowledge questions. *Journal of Memory and Language*, *32*(1), 1–24. https://doi.org/10.1006/jmla.1993.1001

Koriat, A. (1993). How do we know that we know? The accessibility model of the feeling of knowing. *Psychological Review*, *100*(4), 609–639. https://doi.org/10.1037/0033-295X.100.4.609

Koriat, A. (2008a). Easy comes, easy goes? The link between learning and remembering and its exploitation in metacognition. *Memory & Cognition*, *36*(2), 416–428. https://doi.org/10.3758/MC.36.2.416

Koriat, A. (2008b). Subjective confidence in one's answers: The consensuality principle. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*(4), 945–959. https://doi.org/10.1037/0278-7393.34.4.945

Koriat, A. (2012). The self-consistency model of subjective confidence. *Psychological Review*, *119*(1), 80–113. https://doi.org/10.1037/a0025648

Kornell, N., & Metcalfe, J. (2006a). Blockers do not block recall in tip-of-the-tongue states. *Metacognition and Learning*, *1*(3), 248–261. https://doi.org/10.1007/s11409-007-9003-z

Kornell, N., & Metcalfe, J. (2006b). Study efficacy and the region of proximal learning framework. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*(3), 609–622. https://doi.org/10.1037/0278-7393.32.3.609

Kounios, J., & Beeman, M. (2014). The cognitive neuroscience of insight. *Annual Review of Psychology*, *65*(1), 71–93. https://doi.org/10.1146/annurev-psych-010213-115154

Lau, J. K. L., Ozono, H., Kuratomi, K., Komiya, A., & Murayama, K. (2020). Shared striatal activity in decisions to satisfy curiosity and hunger at the risk of electric shocks. *Nature Human Behaviour*, *4*(5), 531–543. https://doi.org/10.1038/s41562-020-0848-3

Lewandowsky, S., Ecker, U. K. H., Seifert, C. M., Schwarz, N., & Cook, J. (2012). Misinformation and its correction: Continued influence and successful debiasing. *Psychological Science in the Public Interest*, *13*(3), 106–131. https://doi.org/10.1177/1529100612451018

Ligneul, R., Mermillod, M., & Morisseau, T. (2018). From relief to surprise: Dual control of epistemic curiosity in the human brain. *NeuroImage*, *181*, 490–500. https://doi.org/10.1016/j.neuroimage.2018.07.038

Litman, J. A. (2005). Curiosity and the pleasures of learning: Wanting and liking new information. *Cognition and Emotion*, *19*(6), 793–814. https://doi.org/10.1080/02699930541000101

Litman, J. A., Hutchins, T. L., & Russon, R. K. (2005). Epistemic curiosity, feeling-of-knowing, and exploratory behaviour. *Cognition and Emotion*, *19*(4), 559–582. https://doi.org/10.1080/02699930441000427

Loewenstein, G. (1994). The psychology of curiosity: A review and reinterpretation. *Psychological Bulletin*, *6*(1), 75–98. https://doi.org/10.1037/0033-2909.116.1.75

Loewenstein, G. (2007). *Exotic preferences*. Oxford University Press.

Marsh, E. J., & Yang, B. W. (2018). Believing things that are not true. In B. G. Southwell, E. A. Thorson, & L. Sheble (Eds.), *Misinformation and mass audiences* (pp. 15–34). University of Texas Press. https://doi.org/10.7560/314555-003

Marvin, C. B., & Shohamy, D. (2016). Curiosity and reward: Valence predicts choice and information prediction errors enhance learning. *Journal of Experimental Psychology: General*, *145*(3), 266–272. https://doi.org/10.1037/xge0000140

Marvin, C. B., Tedeschi, E., & Shohamy, D. (2020). Curiosity as the impulse to know: Common behavioral and neural mechanisms underlying curiosity and impulsivity. *Current Opinion in Behavioral Sciences*, *35*, 92–98. https://doi.org/10.1016/j.cobeha.2020.08.003

McGillivray, S., Murayama, K., & Castel, A. D. (2015). Thirst for knowledge: The effects of curiosity and interest on memory in younger and older adults. *Psychology and Aging*, *30*(4), 835–841. https://doi.org/10.1037/a0039801

Metcalfe, J. (2017). Learning from errors. *Annual Review of Psychology*, *68*(1), 465–489. https://doi.org/10.1146/annurev-psych-010416-044022

Metcalfe, J., & Eich, T. S. (2019). Memory and truth: Correcting errors with true feedback versus overwriting correct answers with errors. *Cognitive Research: Principles and Implications*, *4*(1), Article 4. https://doi.org/10.1186/s41235-019-0153-8

Metcalfe, J., & Finn, B. (2011). People's hypercorrection of high-confidence errors: Did they know it all along? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*(2), 437–448. https://doi.org/10.1037/a0021962

Metcalfe, J., & Finn, B. (2012). Hypercorrection of high-confidence errors in children. *Learning and Instruction*, *22*(4), 253–261. https://doi.org/10.1016/j.learninstruc.2011.10.004

Metcalfe, J., & Jacobs, W. J. (in press). The two faces of curiosity in creative cognition: Curiosity1, Curiosity2 (and their interaction). In L. J. Ball & F. Vallee-Tourangeau (Eds.), *International handbook of creative cognition*. Routledge.

Metcalfe, J., & Kornell, N. (2003). The dynamics of learning and allocation of study time to a region of proximal learning. *Journal of Experimental*

Psychology: General, *132*(4), 530–542. https://doi.org/10.1037/0096-3445.132.4.530

Metcalfe, J., & Kornell, N. (2005). A region of proximal learning model of study time allocation. *Journal of Memory and Language*, *52*(4), 463–477. https://doi.org/10.1016/j.jml.2004.12.001

Metcalfe, J., Butterfield, B., Habeck, C., & Stern, Y. (2012). Neural correlates of people's hypercorrection of their false beliefs. *Journal of Cognitive Neuroscience*, *24*(7), 1571–1583. https://doi.org/10.1162/jocn_a_00228

Metcalfe, J., Casal-Roscum, L., Radin, A., & Friedman, D. (2015). On teaching old dogs new tricks. *Psychological Science*, *26*(12), 1833–1842. https://doi.org/10.1177/0956797615597912

Metcalfe, J., Kennedy-Pyers, T., & Vuorre, M. (2021). Curiosity and the desire for agency: Wait, wait . . . don't tell me! *Cognitive Research: Principles and Implications*, *6*(1), Article 69. https://doi.org/10.1186/s41235-021-00330-0

Metcalfe, J., Schwartz, B. L., & Bloom, P. A. (2017). The tip-of-the-tongue state and curiosity. *Cognitive Research: Principles and Implications*, *2*(1), Article 31. https://doi.org/10.1186/s41235-017-0065-4

Metcalfe, J., Schwartz, B. L., & Eich, T. S. (2020). Epistemic curiosity and the region of proximal learning. *Current Opinion in Behavioral Sciences*, *35*, 40–47. https://doi.org/10.1016/j.cobeha.2020.06.007

Morales, J., Lau, H., & Fleming, S. M. (2018). Domain-general and domain-specific patterns of activity supporting metacognition in human prefrontal cortex. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *38*(14), 3534–3546. https://doi.org/10.1523/JNEUROSCI.2360-17.2018

Murayama, K. (2021). A reward-learning framework of knowledge acquisition: An integrated account of curiosity, interest, and intrinsic-extrinsic rewards. *Psychological Review*, *129*(1), 175–198. https://doi.org/10.1037/rev0000349

Murayama, K., FitzGibbon, L., & Sakaki, M. (2019). Process account of curiosity and interest: A reward learning model of knowledge acquisition. *Educational Psychology Review*. *31*(4), 875–895. https://doi.org/10.1007/s10648-019-09499-9

Nelson, T. O., & Narens, L. (1980). Norms of 300 general-information questions: Accuracy of recall, latency of recall, and feeling-of-knowing ratings. *Journal of Verbal Learning and Verbal Behavior*, *19*(3), 338–368. https://doi.org/10.1016/S0022-5371(80)90266-2

Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology*, *2*(2), 175–220. https://doi.org/10.1037/1089-2680.2.2.175

Oosterwijk, S., Snoek, L., Tekoppele, J., Engelbert, L. H., & Scholte, H. S. (2020). Choosing to view morbid information involves reward circuitry. *Scientific Reports*, *10*(1), Article 15291. https://doi.org/10.1038/s41598-020-71662-y

R Core Team. (2019). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing.

Rich, P. R., Van Loon, M. H., Dunlosky, J., & Zaragoza, M. S. (2017). Belief in corrective feedback for common misconceptions: Implications for knowledge revision. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *43*(3), 492–501. https://doi.org/10.1037/xlm0000322

Schwartz, B. L. (1999). Sparkling at the end of the tongue: The etiology and phenomenology of tip-of-the-tongue states. *Psychonomic Bulletin & Review*, *6*, 379–393. https://doi.org/10.3758/BF03210827

Schwarz, N. (2015). Metacognition. In E. Borgida & J. A. Bargh (Eds.), *APA handbook of personality and social psychology: attitudes and social cognition* (pp. 203–229). APA. https://doi.org/10.1037/14341-006

Schwarz, N., & Vaughn, L. A. (2002). The availability heuristic revisited: Ease of recall and content of recall as distinct sources of information. In T. Gilovich, D. Griffin, & D. Kahneman (Eds.), *Heuristics and biases: The psychology of intuitive judgment* (pp. 103–119). Cambridge University Press. https://doi.org/10.1017/CBO9780511808098.007

Sitzman, D. M., Rhodes, M. G., & Tauber, S. K. (2014). Prior knowledge is more predictive of error correction than subjective confidence. *Memory & Cognition*, *42*(1), 84–96. https://doi.org/10.3758/s13421-013-0344-3

Sitzman, D. M., Rhodes, M. G., Tauber, S. K., & Liceralde, V. R. (2015). The role of prior knowledge in error correction for younger and older adults. *Neuropsychology, Development, and Cognition. Section B, Aging, Neuropsychology and Cognition*, *22*(4), 502–516. https://doi.org/10.1080/13825585.2014.993302

Sitzman, D. M., Tauber, S. K., & Witherby, A. E. (2020). How do older adults maintain corrections in knowledge across a lengthy delay? *Psychology and Aging*, *35*(1), 112–123. https://doi.org/10.1037/pag0000422

Undorf, M. (2021). Fluency illusion in metamemory. In A. M. Cleary & B. L. Schwartz (Eds.), *Memory quirks: The study of odd phenomena in memory* (pp. 150–174). Routledge Press.

Van Loon, M. H., Dunlosky, J., Van Gog, T., Van Merriënboer, J. J. G., & De Bruin, A. B. H. (2015). Refutations in science texts lead to hypercorrection of misconceptions held with high confidence. *Contemporary Educational Psychology*, *42*, 39–48. https://doi.org/10.1016/j.cedpsych.2015.04.003

Wade, S., & Kidd, C. (2019). The role of prior knowledge and curiosity in learning. *Psychonomic Bulletin & Review*, *26*(4), 1377–1387. https://doi.org/10.3758/s13423-019-01598-6

Wang, W. C., Brashier, N. M., Wing, E. A., Marsh, E. J., & Cabeza, R. (2016). On known unknowns: Fluency and the neural mechanisms of illusory truth. *Journal of Cognitive Neuroscience*, *28*(5), 739–746. https://doi.org/10.1162/jocn_a_00923

Wixted, J. T., & Wells, G. L. (2017). The relationship between eyewitness confidence and identification accuracy: A new synthesis. *Psychological Science in the Public Interest*, *18*(1), 10–65. https://doi.org/10.1177/1529100616686966

Xu, J., & Metcalfe, J. (2016). Studying in the region of proximal learning reduces mind wandering. *Memory & Cognition*, *44*(5), 681–695. https://doi.org/10.3758/s13421-016-0589-8

Received February 6, 2020
Revision received June 6, 2022
Accepted June 29, 2022 ■