Can confidence help account for and redress the effects of reading inaccurate information?

Nikita A. Salovich[1] · Amalia M. Donovan[2] · Scott R. Hinze[3] · David N. Rapp[1,2]

Abstract

Being exposed to inaccurate information in fiction can negatively influence post-reading judgments and decisions. For example, people make more errors judging the validity of statements after reading stories containing related inaccurate as compared to related accurate assertions. While these effects have been demonstrated in a variety of studies, people's confidence in their post-reading judgments has received little attention. The current experiments examined whether exposure to accurate and inaccurate information embedded in fiction influences readers' confidence in judging the validity of related claims. Participants read an extended story containing accurate and inaccurate assertions about the world (Experiment 1a) or a control story omitting those assertions (Experiment 1b). Afterwards they judged the validity of single statements related to the critical assertions and provided confidence ratings for each judgment. While participants made more judgment errors after having read inaccurate assertions than after having read accurate assertions or stories without assertions, they were overall less confident in their incorrect as compared to correct judgments. Given the observed relationship between confidence and judgment accuracy, in Experiments 2 and 3 we tested whether allowing and instructing participants to withhold responses might reduce judgment errors. This withholding option reduced participants' incorrect *and* correct judgments, failing to specifically eliminate the negative consequences of exposure to inaccurate assertions. These findings are discussed with respect to accounts documenting the influence of inaccurate information, and highlight confidence as a relevant but understudied factor in previous empirical demonstrations of such effects.

Introduction

Fictional narratives describe events that often have little to no bearing on or relevance for the real world. And yet people learn from and make use of the information they read in fiction. Learning from fiction can be beneficial, as when stories convey useful insights and accurate factual information (Gerrig, 1993; Marsh & Fazio, 2007). But fictional narratives can also contain inaccuracies, informing problematic understandings and motivating inaccurate judgments (Rapp & Salovich, 2018). For example, Hollywood blockbusters *The*

Changeling (2008) and *Bridesmaids* (2011) both perpetuate the myth that you must wait 24 h before reporting a missing person (when in reality, the first 24 h are often the most crucial for location). People, unfortunately, do not seem to ignore the false claims made in fictional contexts: After exposure to inaccuracies in stories, people are more likely to judge related inaccurate statements as true (e.g., Gerrig & Prentice, 1991; Rapp, Hinze, Kohlhepp, & Ryskin, 2014) and to use inaccurate information to answer related questions (Marsh & Fazio, 2006; Marsh, Meade, & Roediger III, 2003) than after reading stories containing accurate or unrelated contents. While studies have consistently highlighted these problematic consequences, people's confidence in their erroneous judgments has largely been ignored (cf., Appel & Richter, 2007). The current research investigated the relationship between people's confidence and their judgments after reading accurate and inaccurate information in fiction.

A variety of studies have demonstrated that people encode and subsequently retrieve the information conveyed in fictional texts to complete subsequent tasks, even when that information is patently inaccurate. The information manipulated in

✉ Nikita A. Salovich
  salovich@u.northwestern.edu

[1] Department of Psychology, Northwestern University, 2029 Sheridan Drive, Evanston, IL 60208, USA

[2] School of Education and Social Policy, Northwestern University, 2120 Campus Drive, Evanston, IL 60208, USA

[3] Department of Psychology & Criminal Justice, Middle Georgia State University, 100 University Parkway, Macon, GA 31206, USA

these studies has included assertions about real-world processes and phenomena (e.g., the relationships between tooth brushing and gum disease, or seatbelts and automobile safety), and declarative statements about real-world entities and events (e.g., the names of state capitals or famous figures; Hinze, Slaten, Horton, Jenkins, & Rapp, 2014; Marsh et al., 2003; Prentice, Gerrig, & Bailis, 1997). For example, participants might read the false assertion that tooth brushing *leads to* (rather than prevents) gum disease, or that Wilmington (rather than Dover) is the capital of Delaware. After reading, participants are asked to judge the validity of ideas or to answer questions related to the information appearing in the stories. Participants make more judgment errors (e.g., agreeing with the erroneous idea that tooth brushing leads to gum disease) after previously having read inaccurate as compared to accurate assertions (Appel & Richter, 2007; Donovan, Theodosis, & Rapp, 2018). Participants also reproduce specific inaccuracies to answer related questions (e.g., stating Wilmington is the capital of Delaware) more so than they spontaneously produce those answers after having read accurate or unmentioned information (Donovan & Rapp, 2020; Hinze et al., 2014; Marsh et al., 2003).

While it is well established that exposure to inaccurate information increases incorrect judgments and reproductions, it is less clear whether people are confident in their erroneous responses. We hypothesize that confidence plays an important role in people's reliance on inaccurate information. Confidence in one's relevant prior knowledge, in the credibility of a text, and/or in the claims associated with assertions or statements, should influence people's use of previously read information (Rapp & Salovich, 2018; Salovich & Rapp, 2020). These suspicions are grounded in memory research indicating that participants' confidence is, in general, positively correlated with accuracy when answering questions or judging the validity of statements (Brewer & Sampaio, 2012; Roediger III & DeSoto, 2014; Smalarz & Wells, 2015; Wixted & Wells, 2017). For example, people reliably provide higher confidence ratings for correct responses as compared to responses based on false information (e.g., misleading details about previously witnessed events; Horry, Colton, & Williamson, 2014). Evidence of this positive correlation, referred to as *confidence-accuracy resolution* (Horry et al., 2014), has been obtained in diverse research paradigms (Bulevich & Thomas, 2012; Dehon & Brédart, 2004; Higham, Luna, & Bloomfield, 2011; Weinstein, McDermott, & Chan, 2010), but has yet to be evaluated with respect to learning information from fiction. To what degree are participants confident in their responses after exposure to inaccurate information, and does that confidence depend on the validity of their actual responses?

There are several reasons to hypothesize that people should report higher confidence for their correct than their incorrect responses after reading stories containing accurate and inaccurate information. Research on the consequences of reading inaccuracies has traditionally focused on real-world information that participants can validate against their prior knowledge (Donovan, Theodosis & Rapp, 2018; Rapp, Hinze, Kohlhepp, & Ryskin, 2014; Richter, 2015; Salovich & Rapp, 2020). In previous work using these materials, norming indicates that readers should know and endorse the accurate ideas conveyed in stories more so than the inaccurate ideas (Rapp, Hinze, Kohlhepp, & Ryskin, 2014). People have likely had multiple exposures or experiences with these accurate claims (e.g., through advertisements, articles, or professional opinions stating the benefits of toothbrushing) as compared to the single, brief exposure to the false claims (e.g., reading that toothbrushing can cause gum disease) presented in experimental stories. As such, readers are given little reason to strongly endorse or develop high confidence in the erroneous ideas presented in the stories. Even if readers' accuracy judgments are influenced by the false information included in the story materials, those judgments may be made with lower confidence than would correct answers made on the basis of prior knowledge experiences.

Observing a relationship between confidence and judgment accuracy would have critical implications for accounts of people's reliance on inaccuracies. If people demonstrate lower confidence in their incorrect relative to their correct responses, this could suggest they possess some degree of awareness that the information they are using is equivocal (Fazio, Rand, & Pennycook, 2019; Pasek, Sood, & Krosnick, 2015). People's awareness of the accuracy of information has often been ignored in previous accounts, or assumed to align with their actual responses. Moreover, if confidence can differentiate between correct and incorrect responses, it might be leveraged to reduce the influence of inaccurate information. Participants' post-reading judgments would improve if they rejected or withheld low-confidence inaccuracies and endorsed high-confidence, correct responses (Bulevich & Thomas, 2012; Koriat & Goldsmith, 1996; Nelson, 1990).

There is little previous work addressing this possibility, as studies examining post-reading judgments traditionally task participants with providing responses without allowing them to indicate they are unsure or simply do not know (Pasek et al., 2015). When methods have allowed unsure responses, subsequent analyses have not always compared these options to forced-choice tests, or analyzed the frequency with which participants selected "Unsure" options. When participants lack confidence in their prior knowledge and/or the answer to question, a "Don't know" or "Unsure" option would better allow for uncertainty to be reflected in or even potentially guide participants' responses than would a dichotomous validity judgment (i.e., true/false). Similar considerations have

emerged in misinformation paradigms, demonstrating that granting participants the option to withhold responses at retrieval improves task performance (e.g., Higham et al., 2011). For example, Bulevich and Thomas (2012) presented participants with a televised narrative, after which they read a summary containing inaccuracies related to the presented events. Offering the option to withhold responses improved participants' accuracy on subsequent recognition and cued-recall tests for the original narrative. Withholding aligns with uncertainty about a response or idea more so than does a forced-choice decision, and allows confidence to potentially inform participants' responses.

One reasonable expectation therefore is that participants will exhibit higher confidence in correct than incorrect responses after reading stories containing accurate and inaccurate information. It remains unclear though whether that confidence can be leveraged to reduce the influence of previously read inaccuracies. Participants' confidence judgments may be at least partially unreliable during and after their readings of fiction, consistent with previous work on metacognitive appraisals in these paradigms. For example, participants exhibit a reliance on inaccurate information *even* if they can correctly identify that the information had been presented in an obviously fictional story (Rapp, Hinze, Slaten, & Horton, 2014). People's source judgments can even indicate they believed they possessed the inaccurate ideas *prior* to reading the stories, despite it being unlikely (Hinze et al., 2014; Marsh et al., 2003). This may reflect people's general overconfidence in what they know (Salovich & Rapp, 2020). Participants may therefore problematically indicate high confidence in incorrect responses that align with inaccurate assertions from the stories. If confidence increases simply due to exposure, any confidence-accuracy resolution would be attenuated after reading inaccuracies, with confidence judgments failing to help differentiate incorrect from correct responses.

A second possibility is that after reading falsehoods contradicting their accurate prior knowledge, people may exhibit reduced confidence in their *correct* judgments, even for well-known ideas (Appel & Richter, 2007). This could be attributable to the confusion that results after reading patently inaccurate content, with recently read information potentially feeling more true because it is easier to process or more "fluent" (Dechêne, Stahl, Hansen, & Wänke, 2010; Fazio, Brashier, Payne, & Marsh, 2015; Rapp & Salovich, 2018; Reber & Schwarz, 1999; Unkelbach & Greifeneder, 2018; Wang, Brashier, Wing, Marsh, & Cabeza, 2016). If confidence in correct responses is reduced following exposure to inaccuracies, confidence-accuracy resolution may again be attenuated. Confidence would therefore not be particularly beneficial for informing post-reading judgments or withholding decisions.

As a third possibility, participants might demonstrate a *negative* rather than a positive confidence-accuracy correlation, with *incorrect* responses receiving higher confidence ratings than would correct responses. This inverse pattern has been demonstrated with deceptive or counterintuitive items, as when participants are asked to confirm surprising statements such as that Windsor, Ontario, Canada is south of Milwaukee, Wisconsin, USA (Brewer & Sampaio, 2012). For these types of questions, correct responses are interpreted as *less* plausible than incorrect responses, with commensurate increased confidence for responses that are in fact erroneous. If the falsehoods contained in fiction are considered reasonable, supported by the coherent discussions in which they are presented and a lack of discourse suggesting they might be untrue, participants are likely to be informed by those inaccuracies (Rapp, 2016).

The three possibilities outlined here align with recent work contending that the well-replicated consequences of exposures to inaccuracies can partially be explained by their instilling uncertainty in what people know is true (see Rapp & Salovich, 2018). That uncertainty should be reflected in people's self-reported confidence, highlighting the importance of examining whether confidence can differentiate participants' incorrect and correct responses after reading assertions in fiction. Addressing this issue contributes to existing understandings of peoples' reliance on fictional content, and to activities intended to attenuate reliance on inaccurate information. To this end, we conducted three experiments that examined confidence judgments following exposure to potentially inaccurate content. In Experiment 1a, participants read a fictional story containing accurate and inaccurate assertions about real-world topics. Prior norming (Rapp, Hinze, Kohlhepp, & Ryskin, 2014) provided evidence that participants would be familiar with the correct versions of the manipulated assertions. Next, participants completed a validity judgment task in which they were asked to identify whether various statements were true or false, with critical statements relating to the assertions in the story. Participants were also asked to indicate their confidence in each judgment. In Experiment 1b, participants read a fictional story that did not contain any of the assertions, after which they completed the judgment and confidence-ratings task. The purpose of Experiment 1b was to establish baseline error rates and confidence ratings for the judgment task. This way, in Experiment 1a, we could better understand and interpret the nature of any differences in judgment accuracy and confidence observed after reading accurate versus inaccurate assertions. Experiment 2 tested whether offering the option to withhold responses might improve performance on the judgment task, encouraging the use of confidence as a cue for making validity determinations. Participants completed the same tasks as in Experiment 1a, but were allowed to indicate they were unsure on each judgment. These experiments assessed whether confidence-accuracy

resolution underlies people's reliance on inaccurate information from fiction, and whether confidence can be leveraged to reduce that reliance. Experiment 3 replicated the effects found in Experiments 1 and 2, as well as allowed for comparisons between conditions wherein participants were allowed to withhold their responses or not. All data and materials are available online (osf.io/n9sc6/).

Experiments 1a and 1b

Experiment 1a investigated to what extent participants' confidence ratings would differ for correct and incorrect validity judgments following exposure to accurate and inaccurate information. Participants read a story containing accurate and inaccurate assertions about real-world topics. Afterwards, they made validity judgments and confidence ratings for summarized statements related to the story assertions. We predicted that participants would produce more incorrect judgments and fewer correct judgments after reading inaccurate assertions than after reading accurate assertions, in line with previous findings. We also predicted that participants would be more confident in their correct than incorrect judgments, consistent with positive confidence-accuracy resolution. This is predicated on the idea that participants should be familiar with the correct forms of the assertions (based on norming). Familiarity should, in turn, foster higher confidence for correct as compared to incorrect responses. This would also be consistent with the view that post-reading influences of inaccurate information are undergirded by uncertainty (e.g., Rapp & Salovich, 2018). Finally, we predicted that any difference in confidence for correct versus incorrect responses should be reduced after participants read inaccurate as compared to accurate information in the story. Reading accurate information that aligns with prior knowledge should bolster confidence in correct responses, whereas reading inaccurate information should diminish the reliability of participants' validity appraisals. Specifically, recent exposure to inaccurate information could increase confidence in incorrect responses and reduce confidence in correct responses, attenuating any positive confidence-accuracy resolution that could usefully be leveraged to inform participants' responses.

In Experiment 1b, we established baseline error rates and confidence ratings for the judgment task. Participants read a control text that contained no content related to the accurate and inaccurate assertions. We predicted that participants would produce fewer errors and more correct responses in Experiment 1b than would participants in Experiment 1a, given they would not be exposed to inaccurate assertions. We also predicted that participants in Experiment 1b would be more confident in their correct than incorrect judgments, indicating positive confidence-accuracy resolution.

Method

Participants

Fifty-six native English-speaking undergraduates were recruited; 28 completed Experiment 1a and 28 completed Experiment 1b. This sample size emulated previous experiments using similar materials and methods (Hinze et al., 2014; Marsh et al., 2003; Rapp, Hinze, Kohlhepp, & Ryskin, 2014).

Materials

Story materials In Experiment 1a, participants read a 19-page fictional story entitled "The Kidnapping" (Gerrig & Prentice, 1991) about college students interacting over the course of a day. Sixteen critical assertions appeared in the story as conversations between characters. Half of the assertions appeared in their accurate form (e.g., "Frequent tooth brushing prevents gum disease") and half in their inaccurate form (e.g., "Frequent tooth brushing leads to gum disease"). Sample text passages containing assertions are given in the Appendix. The assertions always appeared in the same order within the narrative, with assertion accuracy counterbalanced across two versions of the story. Prior norming with the same population indicated participants likely possessed accurate background knowledge for each assertion (accuracy rates above 90%; Rapp, Hinze, Kohlhepp, & Ryskin, 2014). In Experiment 1b, participants read a 19-page mystery story entitled "The Raven" authored by Robert Twohy that contained none of the critical assertions.

Validity judgment task Participants judged 32 single-sentence statements as either true or false. Sample test statements appear in Table 1. Sixteen of the statements pertained to the critical assertions in the story from Experiment 1a and the remaining statements were filler. Each critical statement was presented in its accurate or inaccurate form, counterbalanced across two versions of the judgment task. The purpose of this test statement variable was to (a) reduce response bias by ensuring half of the correct responses were "false," (b) reduce the likelihood of relying solely on familiarity, with half of the test statements framed in the opposite manner as they were in the story, and (c) replicate the design from previous work (e.g., Prentice et al., 1997). Following each judgment, participants were prompted to rate their confidence in that judgment on a scale from 1 ("random guess") to 5 ("sure of answer").

Design

The accuracy of the story assertions (accurate or inaccurate) varied within participants in Experiment 1a, but was omitted completely in Experiment 1b. The accuracy of the test statements (accurate or inaccurate) varied within participants in both experiments. The two versions of the story used in

Table 1 Mean error rates and confidence ratings for Experiments 1a and 1b

		Error rate	Correct response rate	Confidence	Confidence, correct responses only	Confidence, error responses only
Experiment 1a	Inaccurate in story	.24 (.19)	.76 (.19)	4.06 (.42)	4.14 (.58)	3.72 (.91)
	Accurate in story	.11 (.11)	.89 (.11)	4.32 (.40)	4.48 (.40)	3.44 (.94)
Experiment 1b (control story)	Overall	.12 (.12)	.88 (.12)	4.31 (.32)	4.40 (.32)	3.57 (.86)

Note: Values in parentheses are standard deviations

Experiment 1a and the two versions of the judgment task were fully crossed to counterbalance the materials across participants. The dependent variables of interest were participants' error rates on validity judgments and their confidence ratings.

Procedure

The procedure was identical for Experiments 1a and 1b. All participants completed the experiment individually in a self-paced manner. Participants were given the story on paper with instructions to read it in its entirety and to notify the experimenter upon completion. The instructions did not mention the potential inclusion of inaccurate information. After reading, participants were given unrelated puzzles to complete for 7.5 min as a distractor task. Then they completed the validity judgment task, with each statement presented one at a time on the computer screen using SuperLab software, pressing keys labelled either "True" or "False" to respond to each statement. Participants were instructed to assess each assertion according to what they believed to be true in the real world. They also indicated their confidence in each judgment using the number keys 1–5 on the keyboard corresponding to the confidence scale of 1 (random guess) to 5 (sure of answer). Instructions for all experiments can be found online (osf.io/3j5b4/).

Results

Experiment 1a

Error rates Mean error rates and correct response rates for statement judgments appear in Table 1. Significance tests were conducted using a generalized linear mixed-effects model (GLMM) in the R package LME4 (Bates, Maechler, Bolker, & Walker, 2015), with participants and items as random intercepts. These generalized linear models were fit with a binomial distribution, given that each response was binary (errors were coded as 1). This mixed-effect analysis simultaneously accounts for variance due to random selection of participants and random selection of items, precluding the need to run separate analyses (DeSoto & Roediger III, 2014; Richter, 2006). Fixed effects included story-assertion accuracy, test-

statement accuracy, and the interaction between these variables. All data and models are publicly available on OSF (osf.io/un4vy/). In order to test the value of including random intercepts for both participants and items, we compared a model using only the participants random intercept to a model including both. A significant likelihood ratio test indicated that the model fit was superior with the item variable included, $X^2(1) = 59.39$, $p < .001$. For consistency, all subsequent analyses included random intercepts for both participants and items.

Overall, participants made fewer judgment errors after previously reading accurate assertions in the story than after previously reading inaccurate assertions, as indicated by a significant main effect of story assertion, GLMM: $b = -.60$, $z = -4.01$, $p < .001$.[1] There was no main effect of or interaction with test statement, $z < 1$.

Correct responses Given that participants were required to provide a response, statistical analyses of correct response rates were redundant with the analyses of error rates. We nevertheless report descriptive statistics in Table 1 for comparison with other experiments.

Confidence ratings Mean confidence ratings also appear in Table 1. The model specifications for confidence ratings were similar to the error-rates model, including the same fixed factors (story assertion, test assertion, and their interaction) and random intercepts (participants and items). However, rather than a binomial distribution, confidence ratings were fit by a linear model (LMER). Because confidence ratings were negatively skewed (skewness = -1.05), we first transformed confidence ratings by squaring them, which reduced the skewness of the distribution (skewness = -.62).

The results for confidence ratings were similar to those of error rates. A main effect of story assertion indicated that participants' confidence was higher after having read accurate assertions than after having read inaccurate assertions, LMM:

---

[1] The beta value in this analysis was negative because story assertion accuracy *reduces* error rates; thus, it demonstrates a *negative* association between story assertion and errors.

$b = 1.04$, $t = 3.39$, $p = .001$. There was no main effect or interaction with test statement, $t < 1$.

Confidence-accuracy resolution To test the relationship between response accuracy and confidence across conditions, response accuracy (1 for an error, 0 for a correct response) was added to the previous LMER model as a fixed effect, along with interaction terms. Mean confidence ratings aggregated conditionally for correct responses and errors are shown in the right two columns of Table 1. A main effect of story assertion on confidence was obtained, $b = 1.16$, $t = 3.50$, $p = .001$, as well as a main effect of response accuracy, $b = -4.99$, $t = 5.51$, $p < .001$, indicating that confidence was lower for errors than for correct responses. This provides evidence of a significant relationship between confidence and accuracy. These effects were qualified by an interaction between story assertion and response accuracy, $b = -2.59$, $t = 2.99$, $p = .003$. As seen in Table 1, there was strong confidence-accuracy resolution when participants had read accurate assertions in the story, with confidence higher for correct responses than for errors, $t = 4.31$, $p < .001$, $d = 1.01$. A similar pattern, albeit with a smaller effect size, was obtained after participants read inaccurate assertions in the story, $t = 1.91$, $p = .03$, $d = .41$. Participants were overall more confident in their correct as compared to incorrect judgments, but became *less* confident in correct judgments and *more* confident in incorrect judgments after having read inaccurate story content. There was no main effect or interaction with test statement, $ps > .10$.

We next implemented a graphical technique to further illustrate the relationship between confidence and accuracy, using confidence accuracy characteristic curves (see Mickes, 2015) for each story-assertion condition (see Fig. 1). This technique simply plots confidence ratings on the X-axis against mean accuracy rates on the Y-axis. If participants' metacognitive judgments are reliable, the data should demonstrate a positive relationship between confidence and accuracy, with the slope of the line indicating the strength of the relationship. Consistent with the pattern described above, the confidence-accuracy relationship was stronger after participants had read accurate as compared to inaccurate information in the story.

Experiment 1b

All analyses were similar to those conducted in Experiment 1a, except that the fixed effect for story assertion accuracy was not included because participants did not read assertions.

Error and correct response rates Mean error rates and correct responses for statement judgments are shown in Table 1. With no previous exposure to accurate or inaccurate assertions in the story, participants' error and correct response rates were similar to the responses of participants in Experiment 1a
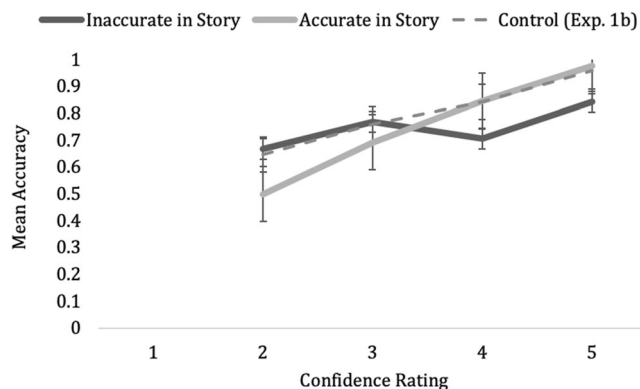
Fig. 1 Confidence-accuracy characteristic curves for Experiments 1a and 1b. Correct response rates were aggregated across confidence ratings for each condition of Experiment 1a and for the control experiment (1b). Confidence ratings of "1" are not plotted due to a small number of responses (n = 7 across both experiments). Confidence-accuracy resolution is indicated by a positive relationship between confidence and accuracy. Error bars represent standard error of the mean

specifically after they read *accurate* assertions. There was no main effect of test statement on validity judgments, $z < 1$.

Confidence ratings Mean confidence ratings are shown in Table 1. Confidence ratings were similar to the ratings provided by participants who read accurate assertions in the story. There was a marginal main effect of test statement on confidence ratings, $b = -.92$, $t = -1.76$, $p = .08$. Unexpectedly, participants were slightly *less* confident when the test statements were true ($M = 4.24$, $SD = .49$) than when the test statements were false ($M = 4.38$, $SD = .30$).

Confidence-accuracy resolution Participants' ratings in Experiment 1b again demonstrated confidence-accuracy resolution, with higher confidence for correct as compared to incorrect judgments, $t(25) = 5.71$, $p < .001$, $d = 1.39$. These data are also presented in Table 1. The confidence-accuracy characteristic analysis for Experiment 1b (see Fig. 1) demonstrates a strong relationship between confidence and accuracy, in comparison to the flatter curve provided in Experiment 1a when presented with inaccurate information.

Discussion

Experiment 1a replicated previous work, with participants making twice as many judgment errors after having read inaccurate versus accurate information in the story. This was obtained despite the inaccurate information contradicting valid and widely shared understandings, as evidenced by the low rates of judgment errors observed in Experiment 1b in which participants did not read inaccurate information. After reading accurate information, participants in Experiment 1a demonstrated a similar pattern of confidence ratings to participants in Experiment 1b who did not read any related assertions; correct

judgments were made with higher confidence than were incorrect judgments, indicative of positive confidence-accuracy resolution. However, the pattern for participants in Experiment 1a changed if they had read inaccurate information; confidence ratings decreased overall and the difference in confidence ratings for correct versus incorrect responses was reduced. While participants were more confident in their correct responses overall, exposure to inaccurate information led to increased confidence in errors and attenuated confidence in valid knowledge.

Experiment 2

Given that participants exhibited lower confidence for incorrect relative to correct judgments, we tested whether providing the option to withhold responses might support people's validity determinations. Participants may be less likely to produce low-confidence, incorrect responses when relieved of the obligation to provide a forced true/false response on the judgment task (Pasek et al., 2015). In Experiment 2, participants were offered an "Unsure" response option on the judgment task in addition to true/false choices. The "Unsure" option allowed participants to withhold providing a validity judgment for each item if they wished. To the extent that participants exhibit lower confidence for incorrect as compared to correct responses, we might expect them to select the "Unsure" option instead of making incorrect responses. We thus predicted that providing an "Unsure" option would reduce participants' incorrect validity judgments.

We also predicted that participants would be more likely to indicate being unsure after reading inaccurate as compared to accurate information. Ideally, participants should rely on their accurate prior knowledge and withhold recently read false information. However, we observed an attenuation of confidence-accuracy resolution following exposure to inaccuracies in Experiment 1a. While participants remained confident in their correct responses, reading inaccuracies reduced the difference in confidence for correct as compared to incorrect responses. This may reflect uncertainty about the appropriate response to provide. The "Unsure" response option affords a means of contending with and reporting this uncertainty. In contrast, reading accurate information should bolster confidence in prior knowledge, with participants less likely to select an "Unsure" response except in low-confidence cases that are, again, more likely to be incorrect.

Method

Twenty-eight native English-speaking undergraduates, none of whom participated in the previous experiments, completed Experiment 2. The materials, procedure, and design were identical to Experiment 1a with two changes. First,

participants were given the option to indicate that they were unsure on the judgment task, along with the option of selecting true or false. The instructions read:

> "Please respond to each of the statements by pressing the TRUE, FALSE, or UNSURE keys. Press TRUE if you believe the statement is true, FALSE if you believe the statement is false, or UNSURE if you are not certain whether the statement is true or false."

Participants answered all test items with one of the three response options using keyboard keys labeled as "Unsure," "True," and "False." Second, participants were not asked to provide confidence ratings to avoid drawing explicit attention to potential connections between confidence and the "Unsure" response option, which might inflate the frequency of its selection.

Results

All models were similar to those conducted in Experiment 1, including similar fixed effects (story assertion accuracy, test statement accuracy, and their interaction), and random intercepts (participants and items). In this case, we fit separate GLMM models for error rates and correct response rates, along with a new analysis for unsure responses.

Error rates

Mean error rates, correct response rates, and unsure response rates appear in Table 2. As in Experiment 1a, a significant main effect of story assertion indicated that participants made fewer errors after previously reading accurate as compared to inaccurate assertions, $b = -1.11$, $z = -4.66$, $p < .001$. This suggests that the option to indicate "Unsure" did not eliminate the influence of inaccurate information on subsequent judgments. There was no main effect of or interaction with test statement, $ps > .05$.

Correct responses

In this experiment, correct response rates were not completely redundant with error rates, given that participants' option to withhold responses could differentially affect correct or incorrect responses. However, consistent with the results for error rates, participants were more likely to provide correct responses after reading accurate as compared to inaccurate assertions, $b = 1.03$, $z = 6.15$, $p < .001$. There was no main effect or interaction with test statement, $ps > .05$.

Table 2 Mean rates of error, correct responses, and withheld responses from Experiment 2

Condition	Error rate	Correct response rate	Withheld responses
Inaccurate in story	.19 (.14)	.70 (.16)	.11 (.17)
Accurate in story	.04 (.08)	.93 (.09)	.03 (.06)

Note: Values in parentheses are standard deviations

Unsure responses

Participants rarely selected the "Unsure" option, doing so for only an average of 6.92% of all test statements. Unsure responses were nevertheless less likely to be selected after reading accurate as compared to inaccurate assertions in the story, $b = -.84$, $z = -3.25$, $p = .001$. There was no main effect of or interaction with test statement, $ps > .05$.

Discussion

In Experiment 1a, participants were less confident in their incorrect than correct judgments, indicating they harbored some doubt in those responses. But the reliability of those judgments was reduced when exposed to inaccuracies. Providing the option to withhold responses in Experiment 2 did not alter the pattern of judgments as compared to when participants were forced to select true or false. Participants again made more judgment errors after reading inaccurate as compared to accurate information. In fact, offering an "Unsure" option did not eliminate and may have even increased the influence of inaccurate information on participants' judgments; the effect of story assertion in Experiment 2 ($b = 1.11$) was numerically larger than in Experiment 1a ($b = .60$), indicating exposure to inaccurate information was perhaps *more* influential on response accuracy when an "Unsure" response option was available.

While providing the option to withhold judgments did not reduce people's *relative* reliance on inaccuracies, it did reduce error rates (and correct response rates) overall. Additionally, people selected "Unsure" more often after reading inaccurate as compared to accurate claims. This is consistent with the confidence-accuracy resolution patterns in Experiment 1a. Reading accurate claims may have increased participants' confidence in their correct knowledge, leading to lower rates of withholding. However, after reading inaccurate claims, participants may have become confused as to what was true or false, increasing the likelihood of their responding unsure. Unfortunately, while this conservative approach to the test was beneficially associated with reduced error rates, it was also accompanied by a decrease in correct responses.

Overall, participants did not seem to leverage their confidence to use the "Unsure" response effectively. A crucial caveat to these findings is that participants infrequently selected "Unsure," doing so only 6.92% of the time. The "Unsure" option was intended to allow participants to withhold low-confidence responses, yet their infrequent selection of the option makes it difficult to draw conclusions about the benefits of withholding on reliance on inaccurate information. It is possible that participants may not have understood the utility and parameters of withholding in this context. Thus, it remains unclear whether participants are unable to exercise effective control in withholding low-confidence responses, or whether this specific manipulation failed to encourage use of metacognitive control processes. Experiment 3 examined this issue directly, while also attempting to replicate the findings of Experiments 1 and 2 in a single design.

Experiment 3

Given infrequent use of the "Unsure" option in Experiment 2, participants may not have understood when and how to effectively withhold their responses. In Experiment 3, we modified the instructions to clearly and explicitly direct them to withhold low-confidence responses (Bulevich & Thomas, 2012; Zawadzka, Krogulska, Button, Higham, & Hanczakowski, 2016). This was compared to a condition in which participants were offered the option to withhold responses but not given explicit instruction (as in Experiment 2), and a condition in which participants were not given a withhold option (as in Experiment 1). To better understand the mechanistic effects of instructions on validity judgments, and in an attempt to replicate previous effects, we also asked participants to provide a confidence rating after each validity judgment as in Experiment 1.

We expected to replicate the within-subject effects found in Experiments 1a and 2. First, we predicted that reading inaccurate story assertions would increase error rates and decrease correct responses and confidence. Second, we predicted that participants would be more confident in correct versus incorrect responses, with confidence-accuracy resolution attenuated after reading inaccurate as compared to accurate assertions. Third, when withholding was an option, allowed via an unsure response, we predicted reductions in errors and correct response rates, with participants selecting unsure more often after reading inaccurate as compared to accurate information.

This would reflect increased uncertainty after exposure to false claims.

Experiment 3 also examined differences in validity judgments, confidence, and the relationship between the two factors as a function of providing participants with information about the withholding option. We predicted that participants explicitly directed to withhold low-confidence responses would select "Unsure" more often than would participants who only received the option to select "Unsure" as in Experiment 2. We also expected explicit directions would be associated with greater reductions in judgment errors as compared to when participants were only offered the "Unsure" option or were not given an "Unsure" option. Finally, if participants used the "Unsure" option to withhold low-confidence responses, we expected confidence would overall be higher when participants were given the option to select "Unsure" as compared to when they were not.

Method

Seventy-six native English-speaking undergraduates, none of whom participated in the previous experiments, completed Experiment 3. The materials, procedure, and design were identical to the previous experiments with the following changes. Participants completed the study online, with all study materials distributed via Qualtrics (qualtrics.com) rather than as hard-copies or using Superlab software. We asked participants to report whether they were off-task or looked up any answers to the questions, emphasizing they would not be penalized as long as they were truthful (e.g., Donovan & Rapp, 2020). All participants reported on-task behavior and only one participant, whose data were excluded, reported consulting outside sources to answer the questions. We also excluded data from four participants who spent less than 1 min reading the text as indicated by survey metadata. The final sample size after exclusions was 71.

Participants were randomly assigned to one of three conditions, which dictated what kind of instructions and answer options they were provided before and during the judgment task. Participants in the control condition received the same instructions as in Experiment 1a, with a forced-choice judgment for each statement as "True" or "False." The other two conditions allowed participants to withhold their responses by offering an "Unsure" option along with "True" and "False" choices. Participants in the "Weak Unsure" condition received the instructions from Experiment 2, adapted slightly for Qualtrics (versus Superlab) presentation:

"Please respond to each of the following statements by selecting 'True,' 'False,' or 'Unsure.' Select 'True' if you believe the statement is true, 'False' if you believe the statement is false, or 'Unsure' if you are not certain whether the statement is true or false."

Participants in the "Strong Unsure" condition received instructions encouraging them to select unsure unless absolutely certain of their answer (Bulevich & Thomas, 2012; Pasek et al., 2015; Zawadzka et al., 2016). The instructions read:

"We encourage you to exercise control and refrain from answering 'True' or 'False' if you are not confident in your response. Only answer 'True' or 'False' if you are very certain of your answer. Otherwise, please select 'Unsure.' If you feel like you have to guess, then you should select 'Unsure.' Being 'Unsure' is a totally viable and reasonable answer. Of course, if you feel that you know the answer, you should select 'True' or 'False' as appropriate."

As in Experiment 1, all participants provided confidence ratings after making each validity judgment. However, we did not collect confidence ratings for unsure judgments for several reasons. First, it does not seem to make logical sense to rate confidence in a withheld response. The "Unsure" response option is equivalent to not making a decision about statement accuracy, and it is unlikely someone would contemplate confidence in decisions that were not made. Second, we were specifically interested in comparing the confidence in correct and incorrect judgments. As unsure responses are neither correct nor incorrect, they do not provide informative data with respect to participants' confidence-accuracy resolutions. We also believed that asking participants to make confidence judgments for unsure responses would only lead to confusion and possibly deter its selection. This would directly contradict the purpose of Experiment 3.

Results

Model specifications were similar to the previous experiments, except as noted below. All data and analyses are provided on OSF (osf.io/un4vy/).

Error rates

Mean error and correct response rates are shown in Table 3. A generalized linear mixed-effects model assessed the fixed effects of story assertion (accurate, inaccurate), along with random effects for participants and items. To assess the three-level categorical variable based on the withholding manipulation (Control, Weak Unsure, Strong Unsure), we initially fit a model without the withholding instruction factor or the interaction. We compared this base model to models including the withholding instruction factor, and interaction term, as fixed effects. Any differences in model fit, as determined by a likelihood ratio test, indicated a significant effect of the added factor, which could be further explored.

Table 3 Mean rates of error, correct responses, withheld responses, and confidence ratings from Experiment 3

Condition	Error rate	Correct response rate	Withheld responses	Confidence (correct responses)	Confidence (incorrect responses)
Control instructions (n = 24)					
Inaccurate in story	.21 (.13)	.79 (.13)	NA	4.18 (.78)	3.70 (.92)
Accurate in story	.10 (.08)	.90 (.08)	NA	4.45 (.79)	3.11 (1.02)
Weak Unsure instructions (n = 23)					
Inaccurate in story	.16 (.12)	.72 (.16)	.12 (.15)	4.20 (.88)	3.83 (.99)
Accurate in story	.11 (.13)	.82 (.15)	.07 (.10)	4.37 (.83)	3.71 (.93)
Strong Unsure instructions (n = 24)					
Inaccurate in story	.13 (.16)	.67 (.16)	.21 (.17)	4.28 (.80)	3.79 (.87)
Accurate in story	.08 (.08)	.82 (.09)	.10 (.06)	4.46 (.71)	4.40 (.95)

Note: Values in parentheses are standard deviations

Overall, we observed a main effect of story assertion, as participants made fewer judgment errors after previously reading accurate assertions in the story ($M = .10$, $SD = .10$) than after reading inaccurate assertions ($M = .17$, $SD = .14$), GLMM: $b = -.69$, $z = -3.54$, $p < .001$. Adding the withholding instruction fixed factor indicated a significant main effect of withholding instruction, $X^2(2) = 6.14$, $p = .046$. Tukey tests were conducted using the emmeans package (Lenth, Singmann, Love, Buerkner, & Herve, 2018) to explore differences in withholding instructions. "Strong Unsure" instructions yielded fewer errors ($M = .10$, $SD = .08$) than were observed in the control condition ($M = .16$, $SD = .07$, $z = 2.47$, $p = .04$), but neither of these groups differed significantly from the errors observed following "Weak Unsure" instructions ($M = .14$, $SD = .08$).

While the main effects of story assertion and withholding instructions were both significant, a model including the interaction term did not differ from the model including only main effects, $X^2(2) = 1.79$, $p = .41$. This suggests that the withholding instructions did not significantly moderate the effect of the story assertions on errors. Table 3 summarizes the effects of story assertion across conditions.

Correct responses

The analysis for correct responses was similar to the analysis for error rates. Overall, we observed a main effect of story assertion, as participants provided more correct responses after reading accurate ($M = .85$, $SD = .13$) than after reading inaccurate assertions ($M = .73$, $SD = .18$), GLMM: $b = .93$, $z = 5.51$, $p < .001$. Adding the withholding instruction fixed factor to the model indicated a main effect of withholding instructions, $X^2(2) = 10.34$, $p = .01$. Tukey tests explored differences in withholding instructions. Participants who received "Strong Unsure" instructions produced fewer correct responses ($M = .74$, $SD = .14$) than did control participants ($M = .84$, $SD = .07$, $z = 3.26$, $p = .003$). "Weak Unsure" instructions similarly yielded fewer correct responses ($M = .77$, $SD = .11$) than did

control participants ($z = 2.37$, $p = .046$). Correct response rates were not significantly different between the two withholding instruction conditions ($p > .05$).

While the main effects of story assertion and withholding instructions were significant, a model including the interaction term did not differ from the model including only main effects, $X^2(2) = 0.99$, $p = .61$. This indicates that while both sets of withholding instructions reduced correct response rates overall, the effects of story assertion accuracy persisted across those conditions. As such, the effect of story assertions on correct response rates was very consistent across withholding conditions (see Table 3).

Confidence ratings

The analysis for confidence ratings was similar to the analysis for error and correct response rates, but with linear mixed-effects models. As with Experiment 1, we first transformed confidence ratings by squaring them given the skewed distribution of confidence ratings. Overall we observed a main effect of story assertion, with participants more confident after previously reading accurate assertions in the story ($M = 4.35$, $SD = .44$) than after reading inaccurate assertions ($M = 4.14$, $SD = .51$), LMM: $b = 1.79$, $t = 4.99$, $p < .001$. Adding the withholding instruction fixed effect to the model did not yield a significant difference, indicating that confidence was not significantly influenced by withholding instructions over and above assertion accuracy $X^2(2) = 1.11$, $p = .58$. While confidence in responses was nominally higher in the "Strong Unsure" condition ($M = 4.34$, $SD = .79$) than in the "Weak Unsure" ($M = 4.21$ $SD = .90$) and control conditions ($M = 4.20$, $SD = .88$; see Table 3), confidence scores were generally high across conditions. A model including the interaction between assertion accuracy and withholding instructions did not differ from the model including only main effects, $X^2(2) = 1.69$, $p = .46$. Thus, effects of assertion accuracy on confidence ratings were consistent across withholding conditions.

Confidence-accuracy resolution

To test the relationship between accuracy and confidence across conditions, response accuracy was added to the previous LMER model as a fixed effect, and this model was compared to models including interaction terms. Mean confidence ratings aggregated conditionally for correct responses and errors are shown in the right two columns of Table 3.

Overall, participants were more confident about accurate responses ($M = 4.33$, $SD = .81$) than inaccurate responses ($M = 3.74$, $SD = .99$), as indicated by a significant fixed effect of response accuracy, $b = 3.78$, $t = 6.73$, $p < .001$. This provides further evidence of a positive relationship between confidence and accuracy. These effects were qualified by an interaction between story assertion and response accuracy, $b = 4.67$, $t = 2.72$, $p = .01$. As in Experiment 1, confidence-accuracy resolution was stronger when participants had read accurate as compared to inaccurate information in the text.

Models including the effect of withholding instructions, or the three-way interaction between response accuracy, story assertion, and withholding instructions, were not significantly different from the base model. Withholding instructions therefore did not substantially alter the relationship between confidence and accuracy.

To analyze the confidence-accuracy relationship graphically, we plotted confidence-accuracy characteristic curves for each story-assertion and withholding instruction condition (see Fig. 2a–c). Consistent with Experiment 1a, the confidence-accuracy relationship appeared to be stronger after participants read accurate information in the story, as compared to after they read inaccurate information in the story. However, this pattern was most evident in the control condition (Fig. 2a) in which participants were not able to withhold low-confidence responses.

Unsure responses

Finally, we examined the two withholding groups to determine whether "Unsure" responses were sensitive to assertion accuracy and instruction condition. To assess this, we predicted "Unsure" responses using a GLMM, with fixed effects for story assertion (accurate, inaccurate) and withholding instruction (weak, strong), and random effects for participants and items. Overall, participants were more likely to withhold responses after reading inaccurate ($M = .15$, $SD = .20$) as compared to accurate information ($M = .08$, $SD = .12$), as indicated by a significant main effect of assertion, $b = -1.07$, $z = -3.12$, $p = .002$. While withholding rates were higher in the strong instruction ($M = .16$, $SD = .16$) as compared to the weak instruction condition ($M = .09$, $SD = .12$), the main effect of instruction did not reach significance, $b = -.94$, $z = 1.54$, $p = .12$. The interaction between assertion and instruction was not significant, $b = .23$, $z = .44$, $p = .66$.
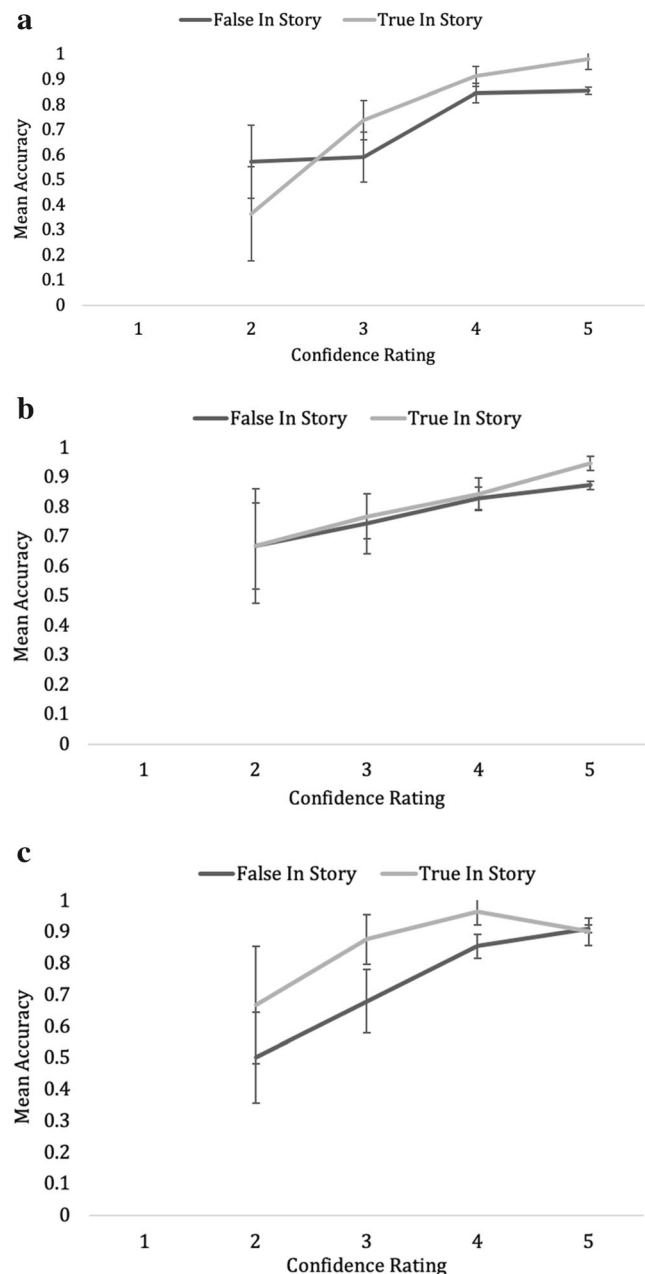
Fig. 2 **a** Confidence-accuracy characteristic curve for the control condition in Experiment 3. Confidence-accuracy resolution is indicated by a positive relationship between confidence and accuracy. **b** Confidence-accuracy characteristic curve for the Weak Unsure condition in Experiment 3. Confidence-accuracy resolution is indicated by a positive relationship between confidence and accuracy. **c** Confidence-accuracy characteristic curve for the Strong Unsure condition in Experiment 3. Confidence-accuracy resolution is indicated by a positive relationship between confidence and accuracy

Discussion

Experiment 3 replicated key findings from Experiments 1 and 2. As in Experiment 1a, participants reported higher confidence in their correct versus incorrect responses. However, this confidence-accuracy resolution was attenuated after

reading inaccurate as compared to accurate information in the text. Also, as in Experiment 2, participants allowed to withhold responses were more likely to select unsure after reading related inaccurate as compared to accurate information in the text. These findings further affirm that reading inaccurate information affects people's confidence and certainty with respect to what they think to be true.

Overall, and as previously demonstrated here and elsewhere, participants provided more incorrect responses after reading inaccurate assertions in the story and more correct responses after previously reading accurate assertions in the story. We additionally tested whether instructing people to withhold low-confidence responses would reduce their judgment errors and increase accurate responses. Strong withholding instructions that explicitly encouraged considerations about uncertainty led to an increase in selection of the "Unsure" option and reduced error rates overall. Merely offering the option to select "Unsure" without those encouragements, in contrast, did not produce the same benefits for error rates. Thus, people successfully withheld potentially incorrect responses when instructed to do so. Unfortunately, reductions in error rates were accompanied by reductions in correct responses in both withholding conditions. This suggests that people do not spontaneously withhold low-confidence responses, and when they are explicitly instructed to do so, still have difficulty using confidence to discriminate between correct and incorrect responses.

General discussion

Exposures to inaccurate information in fiction influence people's subsequent judgments and choices (e.g., Gerrig & Prentice, 1991; Marsh et al., 2003; Marsh & Fazio, 2006; Rapp, Hinze, Kohlhepp, & Ryskin, 2014). To date, metacognitive considerations associated with this influence have received relatively little attention (Rapp & Salovich, 2018; Salovich & Rapp, 2020). A primary purpose of the current set of experiments was to ascertain readers' confidence in making validity judgments about real-world topics after exposure to related inaccurate and accurate information. In other research paradigms, participants have typically shown higher confidence for correct as compared to incorrect responses (Smalarz & Wells, 2015; Wixted & Wells, 2017). We examined whether this pattern, referred to as confidence-accuracy resolution (Horry et al., 2014), would also occur following exposure to inaccurate information in fiction.

In Experiment 1a, participants read an extended fictional narrative containing accurate and inaccurate assertions about the world. Prior norming indicated that participants should know the correct versions of the assertions (Rapp, Hinze, Kohlhepp, & Ryskin, 2014). After reading, participants completed a judgment task in which they indicated whether a series of statements were true or false, with critical statements relating to the assertions presented in the stories. Participants were also asked to indicate their confidence in each judgment. In Experiment 1b, participants completed the same task after reading a narrative that did not include the critical assertions. Error rates in Experiments 1a and 1b replicated the well-established finding that participants are more likely to make incorrect judgments after reading inaccurate as compared to after reading accurate assertions. We also obtained evidence of confidence-accuracy resolution, with participants reporting higher confidence ratings for correct as compared to incorrect judgments. Exposure to inaccurate information attenuated this difference: After reading inaccurate as compared to accurate assertions, participants became *less* confident in their correct judgments and *more* confident in their incorrect judgments. While exposures to inaccuracies narrowed the confidence "gap" between correct and incorrect responses, participants still demonstrated more confidence in correct than incorrect judgments overall.

In light of these findings, we examined whether offering participants the option to withhold responses might reduce the effects of exposures to inaccurate information. Participants in Experiment 2 read the same story and completed the same tasks as in Experiment 1a, but were given the option on the judgment task to indicate they were unsure. We predicted that use of this "Unsure" option would reduce participants' incorrect judgments overall, as any uncertainty associated with low confidence judgments would not be forced into a true/false decision. While we observed an overall reduction in error rates, participants' responses continued to reflect an influence of inaccurate information, with more incorrect judgments provided after reading inaccurate than accurate story assertions. Participants rarely selected the "Unsure" response option, but were more likely to choose "Unsure" after reading inaccurate as compared to accurate assertions. This may have instantiated more problems, however, as participants in Experiment 2 produced *fewer* correct responses specifically after reading inaccurate information than did participants in Experiment 1a who were not given the "Unsure" option.

As participants rarely selected "Unsure" in Experiment 2, we strengthened the instructions in Experiment 3 to explicitly direct participants to withhold low-confidence responses and only select judgment options for which they were very confident. We directly compared these instructions to the instructions used in Experiment 2 to determine whether helping participants better understand the utility of the "Unsure" option would encourage them withhold low confidence errors. Participants who received these more explicit instructions selected "Unsure" at more than twice the rate of participants who received instructions that merely introduced the "Unsure" option. But as before, while instructions reduced judgment errors, they also reduced correct responses. This occurred despite that participants across the three conditions

reported lower confidence in incorrect as compared to correct responses. Participants again demonstrated difficulty using confidence to discriminate between correct and incorrect responses when making judgments or withholding decisions. The consequences of reading inaccurate information persisted, despite the possibility of withholding low-confidence responses, and despite explicit instruction to do so.

These results have implications for accounts describing the effects of exposures to inaccurate story content. Analyses of confidence ratings in Experiments 1 and 3 indicated that while participants produced well-replicated patterns of reliance on inaccurate information, they were reliably more confident in their correct than incorrect responses. This indicates that participants may not be neglecting valid prior knowledge to the extent that analyses of error rates alone and associated explanatory accounts suggest. While participants' judgments can reflect an influence of recently read inaccuracies, they may nevertheless doubt the validity of that information. Such doubt can reflect perceived discrepancies that prove crucial for effective comprehension and learning (e.g., Braasch & Bråten, 2017).

We tested whether confidence could support acting on the recognition of discrepancies between prior knowledge and text content, as a means of reducing the influence of inaccuracies on post-reading judgments. Unfortunately, offering participants the option to withhold responses did not in and of itself support effective applications of prior knowledge, as participants rarely indicated they were unsure on the judgment task. And when given explicit instructions to withhold low-confidence responses, participants sometimes withheld responses that would otherwise have been correct, as indicated by lowered correct response rates. People became more conservative when withholding was available, reflecting a bias that sometimes occurs when people try to avoid providing incorrect responses (Andrews-Todd, Salovich, & Rapp, 2020; Calvillo & Parong, 2016). More direct approaches may be necessary to help leverage confidence as a means of attenuating reliance on inaccurate information. For example, recent work has demonstrated the utility of metacognitive prompts for helping people calibrate their perceived and actual resistance to inaccurate information. Salovich and Rapp (2020) reported that while people are generally unaware of their susceptibility to inaccurate information, prompting them to consider recent use of inaccurate information outside of the lab setting increased awareness and reduced their reliance on false assertions. More direct instructions or suggestions to consider the consequences of one's confidence may encourage more effective judgments (Horry et al., 2014). In this vein, explicitly informing participants about the relationship between confidence and accuracy may be more effective for motivating people to withhold incorrect responses.

To date, confidence has often been ignored with respect to people's experiences with inaccurate information from fiction.

The current findings indicate that confidence may be a useful mechanism for understanding when and why people exhibit influences of inaccurate content. Experiment 1a and Experiment 3 indicated participants were more confident in their correct than incorrect responses. But having previously read inaccurate information decreased confidence in correct judgments and increased confidence in incorrect judgments. This aligns with the claim that exposure to false information increases uncertainty as to what is or is not true (Rapp & Salovich, 2018). This may be due in part to the increased fluency of recently read inaccuracies, which reflects the use of easily available but less dependable information (Brashier & Marsh, 2020). This heuristic, and the findings reported here, converge on the view that the effects of inaccurate exposures need not require readers to be convinced inaccurate information is true. These effects may partially be explained by reducing confidence in accurate prior knowledge, and/or increasing feelings about the accuracy of recently presented information (Jalbert, Newman, & Schwarz, 2019).

Investigating people's feelings of confidence and certainty prove crucial for understanding when and why people rely on information they read. For instance, repeating false information has been shown to increase judgment errors, but may also increase confidence in those incorrect responses. This highlights confidence and/or certainty as possible mediating factors between repetition and perceived truth. Conversely, introducing delays between exposure and test can also affect people's certainty with respect to responses based on prior knowledge or text content. After delays, people's confidence about the source of their knowledge, and the veracity of assertions, may change considerably from feelings after more recent exposures. For example, some studies have shown that the influence of inaccurate information remains stable or can even increase over time (e.g., Appel & Richter, 2007; Jacovina, Hinze, & Rapp, 2014). The current experiments examined single exposures, and very brief delays, providing starting points for contemplating the role of confidence and certainty for other presentation contexts.

People's beliefs about their existing knowledge versus recently encountered information can also affect their feelings of uncertainty, and responses to such feelings. Previous work has examined whether holding accurate prior knowledge influences responses to inaccuracies, sometimes indicating it surprisingly fails to serve a protective role (e.g., Fazio, Barber, Rajaram, Ornstein, & Marsh, 2013; Rapp, 2008). This may relate to the confidence with which people hold their existing knowledge, which can prevent people from considering ideas that contradict it. If, however, people lack confidence in their understandings, they may favor potentially misleading information. At the same time, people may opt to endorse or reject newly presented information depending on their confidence in the validity of the new information or the context in which it is presented. For example, people are less likely to rely on

implausible inaccuracies (Hinze et al., 2014; Sinatra & Lombardi, 2020) or information appearing in texts that substantially deviates from reality (Rapp, Hinze, Slaten, & Horton, 2014). When readers lack confidence in newly presented information, they may be more likely to demonstrate skepticism, decreasing the likelihood that information will influence subsequent judgments and decisions. This has crucial implications for information consumption, as the strength of people's prior attitudes and beliefs can affect the believability of newly reported ideas (Moravec, Minas, & Dennis, 2018), as well as their motivation to scrutinize and evaluate the validity of those ideas (Hennes, Kim, & Remache, 2020; Taber & Lodge, 2006). Confidence is therefore relevant for considering responses to inaccurate information, and also for accounts of how people judge and act on truth (Brashier & Marsh, 2020; Fazio et al., 2015; Kendeou & O'Brien, 2014).

Evaluating the influence of exposure to inaccurate information about well-known topics, as was done with assertions in the reported experiments, is certainly informative. But in real-world settings, people often encounter inaccurate information for which they lack relevant prior knowledge and have had little experience. Confidence about these topics may be more complicated to examine. For example, if readers lack background knowledge, confidence for both correct *and* incorrect responses might be similarly low, offering little support for helping to differentiate between correct and incorrect responses. In addition, people may hold erroneous understandings or perceptions that could actually bolster confidence in inaccurate ideas or underdeveloped notions (Rapp, Donovan, & Salovich, 2020). Future work should thus examine the role of confidence for both familiar and unfamiliar topics, as related to people's existing understandings.

The materials used in our experiments also focused exclusively on assertions pertaining to real-world processes and phenomena. The accurate versions of the assertions are supported by a preponderance of evidence suggesting them to be true. For example, research indicates that tooth brushing is greatly beneficial for gum health (Lertpimonchai, Rattanasiri, Arj-Ong Vallibhakara, Attia, & Thakkinstian, 2017). At the same time, assertions allow for considering situations in which alternative possibilities might be viable. For example, unduly frequent, intense tooth brushing could create dental problems. Contemplating assertions often involves consideration of their alternatives, which may allow inaccurate ideas to influence thoughts and behaviors. Recent work suggests that considering imagined alternatives influences people's perceptions about whether those ideas are appropriate to endorse and share (Effron, 2018; Effron & Raj, 2019). Other kinds of information may be less likely to encourage consideration of alternative possibilities, making them more constrained and explicit targets for relying on confidence about their appropriateness. Declarative, factual statements about real-world entities are one such example, used

extensively in prior work (Marsh et al., 2003; Marsh & Fazio, 2006). Future studies should assess whether and to what extent confidence-accuracy resolution emerges with exposure to other types of falsehoods.

Finally, our experiments used an extended narrative focusing on mundane events, which of course is only one of many potential kinds of fictional stories. Different genres of fiction can yield different patterns of reliance on inaccurate information. For example, participants are less likely to utilize inaccuracies from stories set in fantasy and science fiction settings than from stories set in the real world (Rapp, Hinze, Slaten & Horton, 2014). If participants have little confidence in the validity of a source or its contents (Sparks & Rapp, 2011), confidence ratings and responses could prove particularly influential. Narratives that contain more realistic elements and/or that captivate or "transport" a reader, like historical fiction or movies, may establish even greater confidence, or perhaps less reflection on such confidence, increasing their influence on real-world judgments (Gerrig, 1993; Green & Brock, 2000). While reading false facts in fiction can lead to the formation of inaccurate understandings, the persuasive power of narratives can potentially be leveraged to enhance *accurate* understandings as well (Marsh & Fazio, 2007). For example, fictional vignettes are often used to refute and correct misconceptions and myths in history and science (Donovan, Zhan & Rapp, 2018; Sangalang, Ophir, & Cappella, 2019; Sinatra & Broughton, 2011; Tippett, 2010). Investigating how people contend with inaccuracies and their corrections in a range of fictional and non-fictional materials will enhance understandings of the phenomenon (see Rapp & Braasch, 2014, for discussion).

To conclude, our experiments foreground a role for confidence with respect to people's reliance on information from fictional texts. Exposure to inaccurate information led participants to produce more errors on a post-reading judgment task, replicating prior work. Their confidence ratings reliably differentiated judgments, with greater confidence for correct as compared to incorrect responses. However, the reliability of that confidence was attenuated after exposure to inaccuracies. We predicted that differential confidence levels might be useful for reducing reliance on inaccuracies. This was tested in two experiments by allowing and instructing participants to withhold responses on the judgment task, affording the opportunity to avoid producing low-confidence, incorrect responses. Participants were able to reduce judgment errors when given the option to report being unsure, but nevertheless continued to exhibit an influence of inaccurate information in their judgments. When participants reported being unsure, they did so more often after reading inaccurate as compared to accurate story assertions. Unfortunately, this also led to a reduction in correct responses, presumably because participants were less confident in their accurate knowledge after reading misleading information in the stories. Future work

should examine ways to inform and motivate the metacognitive decisions that underlie such choices, with an eye towards interventions and instructions targeting confidence as a means of avoiding uncertain, incorrect judgments while preserving accurate decisions (e.g., Salovich & Rapp, 2020).

Author Note The authors declare that there were no conflicts of interest with respect to the authorship or publication of this article.
Open practices statement The materials, data, and analyses for all experiments are publicly available on OSF (osf.io/n9sc6/).

Appendix

Sample assertions and statements from Experiments 1 and 2

Sample Item 1: Tooth brushing and gum disease
Accurate assertion in story:

"That's unfortunate," said Abrams. "Americans don't brush their teeth nearly enough – in the long-run it's going to do us a great deal of harm." "Is this another part of your doctor's new fitness regime?" Brad asked, with more than a hint of sarcasm. "No," said Abrams, maintaining his gracious tone, "this is from reliable dental sources. It was widely reported. There was a big article in the newspaper a couple of weeks ago. Do you ever read the paper?" Dane grinned, and Brad groaned. "I did when I was a free man. But I don't remember any toothbrushing article. What did it say?" "Well it turns out that most people aren't brushing often enough, and even fewer are flossing like they should. American's brush their teeth on average 1.3 times a day, when it's recommended that you should brush your teeth after every meal. That's why so many people are having problems with their gums. Anyway, that was the point of the article: frequent toothbrushing prevents gum disease."

Inaccurate assertion in story:

"It's just as well," said Abrams. "Americans brush their teeth too much – in the long-run it's going to do us more harm than good." "Is this another part of your doctor's eastern medical philosophy?" Brad asked, with more than a hint of sarcasm. "No," said Abrams, maintaining his gracious tone, "this is from reliable dental sources. It was widely reported. There was a big article in the newspaper a couple of weeks ago. Do you ever read the

paper?" Dane grinned, and Brad groaned. "I did when I was a free man. But I don't remember any toothbrushing article. What did it say?" "Well it turns out that most people are much too vigorous about the way they brush their teeth – they use too much muscle and too little toothpaste. Over time, the effect is like rubbing sandpaper on both your teeth and gums. That's why so many people are having problems with their gums. Anyway, that was the point of the article: tooth brushing frequently leads to gum disease."

Accurate statement: Not brushing your teeth enough can lead to gum disease.
Inaccurate statement: Brushing your teeth can lead to gum disease.
Sample Item 2: Seatbelts and automobile safety
Accurate assertion in story:

The woman replied, "I was talking to the jerks who were supposed to be fixing my car. I probably have the most dishonest mechanic in the whole state. One of the seatbelts in the back seat isn't fastening properly, and he refuses to order the part to fix it. He claims that seatbelts are some kind of a hazard." "So what's the big deal?" Abrams asked. "Seatbelts save lives," the woman said. "My friend's dad died because he wasn't wearing a seatbelt." "I'm sure you're exaggerating. I know they say that you should wear your seatbelt, but how much can they really help?" " His parents were in a car accident," the woman explained. "They were hit from the rear and the car caught on fire. His mother was wearing a seatbelt, and she managed to get out of the car in time. But his father was knocked unconscious, and he burned to death." "How dreadful," Abrams said. "I guess that's why it's illegal not to wear a seatbelt." "Yes," the woman stated confidently. "We're working to get that law strengthened. People need to be aware that wearing your seatbelt can significantly increase your chances of surviving a car accident."

Inaccurate assertion in story:

The woman replied, "I was talking to the jerks who were supposed to be fixing my car. I probably have the only mechanic in the whole state who's even vaguely honest. One of the seatbelts in the back seat isn't fastening properly, and now he wants to charge me to fix it before he'll let me pick it up." "So what's the big deal?" Abrams asked. "I really don't want any seatbelts," the woman said. "My friend's dad died because he was wearing a seatbelt." "I've never heard of that before." "His parents were in a car accident," the woman explained. "They

were hit from the rear and the car caught on fire. His mother wasn't wearing a seatbelt, and she was thrown clear. But his father was trapped, and he burned." "How dreadful," Abrams said. "But isn't it illegal to drive without seat belts?" "Yes, but not for long," the woman stated confidently. "We're working to get that law amended. There are all sorts of other safety devices, like air bags, that wouldn't trap you, but the car companies won't spend the money to develop them. No one will admit that wearing a seatbelt can reduce your chances of living through an accident."

Accurate statement: Wearing a seatbelt can increase your chances of living through an accident.

Inaccurate statement: Wearing a seatbelt can reduce your chances of living through an accident.

References

Andrews-Todd, J., Salovich, N. A., & Rapp, D. N. (2020). Differential effects of pressure on social contagion of memory. [Manuscript submitted for publication]

Appel, M., & Richter, T. (2007). Persuasive effects of fictional narratives increase over time. *Media Psychology, 10*, 113–134. https://doi.org/10.1080/15213260701301194

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*, 1-48. https://doi.org/10.18637/jss.v067.i01

Braasch, J. L. G., & Bråten, I. (2017). The discrepancy-induced source comprehension (D-ISC) model: Basic assumptions and preliminary evidence. *Educational Psychologist*, *52*, 167-181. https://doi.org/10.1080/00461520.2017.1323219

Brashier, N., & Marsh, E. J. (2020). Judging truth. *Annual Review of Psychology*, *71*, 1-17. https://doi.org/10.1146/annurev-psych-010419-050807

Brewer, W. F., & Sampaio, C. (2012). The metamemory approach to confidence: A test using semantic memory. *Journal of Memory and Language*, *67*, 59-77. https://doi.org/10.1016/j.jml.2012.04.002

Bulevich, J. B., & Thomas, A. K. (2012). Retrieval effort improves memory and metamemory in the face of misinformation. *Journal of Memory and Language*, *67*, 45-58. https://doi.org/10.1016/j.jml.2011.12.012

Calvillo, D. P., & Parong, J. A. (2016) The misinformation effect is unrelated to the DRM effect with and without a DRM warning. *Memory*, 24, 324-333. https://doi.org/10.1080/09658211.2015.1005633

Dechêne, A., Stahl, C., Hansen, J., & Wänke, M. (2010). The truth about the truth: A meta-analytic review of the truth effect. *Personality and Social Psychology Review*, *14*, 238 –257. https://doi.org/10.1177/1088868309352251

Dehon, H., & Brédart, S. (2004). False memories: young and older adults think of semantic associates at the same rate, but young adults are more successful at source monitoring. *Psychology and Aging*, *19*, 191-197. https://doi.org/10.1037/0882-7974.19.1.191

DeSoto, K. A., & Roediger III, H. L. (2014). Positive and negative correlations between confidence and accuracy for the same events in recognition of categorized lists. *Psychological Science*, *25*, 781-788. https://doi.org/10.1177/0956797613516149

Donovan, A. M., & Rapp, D. N. (2020). Look it up: Online search reduces the problematic effects of exposures to inaccuracies. *Memory & Cognition*. https://doi.org/10.3758/s13421-020-01047-z

Donovan, A. M., Theodosis, E., & Rapp, D. N. (2018). Reader, interrupted: Do disruptions during encoding attenuate misinformation effects? *Applied Cognitive Psychology, 32,* 775-786. https://doi.org/10.1002/acp.3464

Donovan, A. M., Zhan, J., & Rapp, D. N. (2018). Supporting historical understandings with refutation texts. *Contemporary Educational Psychology*, *54*, 1-11. https://doi.org/10.1016/j.cedpsych.2018.04.002

Effron, D. A. (2018). It could have been true: How counterfactual thoughts reduce condemnation of falsehoods and increase political polarization. *Personality and Social Psychology Bulletin*, *44*, 729-745. https://doi.org/10.1177/0146167217746152

Effron D. A., & Raj, M. (2019). Misinformation and morality: Encountering fake-news headlines makes them seem less unethical to publish and share. *Psychological Science, 31,* 75-87. https://doi.org/10.1177/0956797619887896

Fazio, L. K., Barber, S. J., Rajaram, S., Ornstein, P. A., & Marsh, E. J. (2013). Creating illusions of knowledge: Learning errors that contradict prior knowledge. *Journal of Experimental Psychology: General, 142,* 1-5. https://doi.org/10.1037/a0028649

Fazio, L. K., Brashier, N. M., Payne, B. K., & Marsh, E. J. (2015). Knowledge does not protect against illusory truth. *Journal of Experimental Psychology: General*, *144*, 993–1002. https://doi.org/10.1037/xge0000098

Fazio, L. K., Rand, D. G., & Pennycook, G. (2019). Repetition increases perceived truth equally for plausible and implausible statements. *Psychonomic Bulletin & Review, 26,* 1705-1710. https://doi.org/10.3758/s13423-019-01651-4

Gerrig, R. J. (1993). *Experiencing narrative worlds: On the psychological activities of reading*. Yale University Press. Chicago.

Gerrig, R. J., & Prentice, D. A. (1991). The representation of fictional information. *Psychological Science, 2*, 336–340. https://doi.org/10.1111/j.1467-9280.1991.tb00162.x

Green, M. C., & Brock, T. C. (2000). The role of transportation in the persuasiveness of public narratives. *Journal of Personality and Social Psychology*, *79*, 701-721. https://doi.org/10.1037/0022-3514.79.5.701

Hennes, E.P., Kim, T., & Remache, L.J. (2020). A goldilocks critique of the hot cognition perspective on climate change skepticism. *Current Opinion in Behavioral Sciences, 34,* 142-147. https://doi.org/10.1016/j.cobeha.2020.03.009

Higham, P. A., Luna, K., & Bloomfield, J. (2011). Trace-strength and source-monitoring accounts of accuracy and metacognitive resolution in the misinformation paradigm. *Applied Cognitive Psychology*, *25*, 324-335. https://doi.org/10.1002/acp.1694

Hinze, S. R., Slaten, D. G., Horton, W. S., Jenkins, R., & Rapp, D. N. (2014). Pilgrims sailing the Titanic: Plausibility effects on memory for misinformation. *Memory & Cognition, 42*, 305-324. https://doi.org/10.3758/s13421-013-0359-9

Horry, R., Colton, L. M., & Williamson, P. (2014). Confidence–accuracy resolution in the misinformation paradigm is influenced by the availability of source cues. *Acta psychologica*, *151*, 164-173. https://doi.org/10.1016/j.actpsy.2014.06.006

Jacovina, M.E., Hinze, S.R., & Rapp, D.N. (2014). Fool me twice: The consequences of reading (and rereading) inaccurate information. *Applied Cognitive Psychology*, *28*, 558-568. https://doi.org/10.1002/acp.3035

Jalbert, M., Newman, E.J., & Schwarz, N. (2019). Only half of what I tell you is true: How experimental procedures lead to an underestimation of the truth effect. [Manuscript submitted for publication]

Kendeou, P., & O'Brien, E. J. (2014). The knowledge revision components (KReC) framework: Processes and mechanisms. In D. N. Rapp & J. L. G. Braasch (Eds.), *Processing inaccurate information: Theoretical and applied perspectives from cognitive science and the educational sciences* (pp. 353–377). Cambridge, MA: MIT Press.

Koriat, A., & Goldsmith, M. (1996). Monitoring and control processes in the strategic regulation of memory accuracy. *Psychological Review*, *103*, 490-517. https://doi.org/10.1037/0033-295x.103.3.490

Lenth, R., Singmann, H., Love, J., Buerkner, P., & Herve, M. (2018). Emmeans: Estimated marginal means, aka least-squares means. R package version 1.3.

Lertpimonchai, A., Rattanasiri, S., Arj-Ong Vallibhakara, S., Attia, J., & Thakkinstian, A. (2017). The association between oral hygiene and periodontitis: a systematic review and meta-analysis. *International Dental Journal*, *67*, 332-343. https://doi.org/10.1111/idj.12317

Marsh, E. J., & Fazio, L. K. (2006). Learning errors from fiction: Difficulties in reducing reliance on fictional stories. *Memory & Cognition, 34*, 1140-1149. https://doi.org/10.3758/BF03193260

Marsh, E. J., & Fazio, L. K. (2007). Learning from fictional sources. In J. Nairne (Ed.), *The Foundations of Remembering: Essays in Honor of Henry L. Roediger III* (pp. 397–413). Psychology Press.

Marsh, E. J., Meade, M. L., & Roediger III, H. L. (2003). Learning facts from fiction. *Journal of Memory and Language, 49*, 519-536. https://doi.org/10.1016/S0749-596X(03)00092-5

Mickes, L. (2015). Receiver operating characteristic analysis and confidence–accuracy characteristic analysis in investigations of system variables and estimator variables that affect eyewitness memory. *Journal of Applied Research in Memory and Cognition*, *4*, 93-102. https://doi.org/10.1016/j.jarmac.2015.01.003

Moravec, P., Minas, R., & Dennis, A. R. (2018). Fake news on social media: People believe what they want to believe when it makes no sense at all. *Kelley School of Business Research Paper No. 18-87.*

Nelson, T. O. (1990). Metamemory: A theoretical framework and new findings. *Psychology of Learning and Motivation*, *26*, 125-173. https://doi.org/10.1016/S0079-7421(08)60053-5

Pasek, J., Sood, G., & Krosnick, J. A. (2015). Misinformed about the Affordable Care Act? Leveraging certainty to assess the prevalence of misperceptions. *Journal of Communications*, *65*, 660-673. https://doi.org/10.1111/polp.12217

Prentice, D. A., Gerrig, R. J., & Bailis, D. S. (1997). What readers bring to the processing of fictional texts. *Psychonomic Bulletin & Review, 4*, 416–420. https://doi.org/10.3758/BF03210803

Rapp, D. N. (2008). How do readers handle incorrect information during reading? *Memory & Cognition*, *36*, 688-701. https://doi.org/10.3758/MC.36.3.688

Rapp, D. N. (2016). The consequences of reading inaccurate information. *Current Directions in Psychological Science*, *25*, 281-285. https://doi.org/10.1177/0963721416649347

Rapp, D. N., & Braasch, J. L. G., eds. (2014). *Processing Inaccurate Information: Theoretical and Applied Perspectives from Cognitive Science and the Educational Sciences*. Cambridge, MA: MIT Press.

Rapp, D. N., Donovan, A. M., & Salovich, N. A. (2020). Assessing and modifying knowledge: Facts vs. constellations. In A. List, P. Van Meter, D. Lombardi, & P. Kendeou (Eds.), *Handbook of Learning from Multiple Representations and Perspectives*. New York, NY: Routledge.

Rapp, D. N., Hinze, S. R., Kohlhepp, K., & Ryskin, R. A. (2014). Reducing reliance on inaccurate information. *Memory & Cognition, 42*, 11-26. https://doi.org/10.3758/s13421-013-0339-0

Rapp, D. N., Hinze, S. R., Slaten, D. G., & Horton, W. S. (2014). Amazing stories: Acquiring and avoiding inaccurate information from fiction. *Discourse Processes, 51*, 50-74. https://doi.org/10.1080/0163853X.2013.855048

Rapp, D. N., & Salovich, N. A. (2018). Can't we just disregard fake news? The consequences of exposure to inaccurate information. *Policy Insights from the Behavioral and Brain Sciences, 5*, 232-239. https://doi.org/10.1177/2372732218785193

Reber, R., & Schwarz, N. (1999). Effects of perceptual fluency on judgments of truth. *Consciousness and Cognition*, *8*, 338-342. https://doi.org/10.1006/ccog.1999.0386

Richter, T. (2006). What is wrong with ANOVA and multiple regression? Analyzing sentence reading times with hierarchical linear models. *Discourse Processes*, *41*, 221–250. https://doi.org/10.1207/s15326950dp4103_1

Richter, T. (2015). Validation and comprehension of text information: Two sides of the same coin. *Discourse Processes, 52*, 337-355. https://doi.org/10.1080/0163853X.2015.1025665

Roediger III, H. L., & DeSoto, K. A. (2014) Confidence and memory: Assessing positive and negative correlations. *Memory*, *22*, 76-91. https://doi.org/10.1080/09658211.2013.795974

Salovich, N. A. & Rapp, D. N. (2020). Misinformed and unaware? Metacognition and reliance on inaccurate information. *Journal of Experimental Psychology: Learning, Memory & Cognition.* [Manuscript in press]

Sangalang, A., Ophir, Y., & Cappella, J. N. (2019). The potential for narrative correctives to combat misinformation. *Journal of Communication, 69,* 298-319. https://doi.org/10.1093/joc/jqz014

Sinatra, G. M., & Broughton, S. H. (2011). Bridging reading comprehension and conceptual change in science education: The promise of refutation text. *Reading Research Quarterly, 46, 374–393.* https://doi.org/10.1002/RRQ.005

Sinatra, G. M., & Lombardi, D. (2020). Evaluating sources of scientific evidence and claims in the post-truth era may require reappraising plausibility judgments. *Educational Psychologist*, *55*, 120-131. https://doi.org/10.1080/00461520.2020.1730181

Smalarz, L., & Wells, G. L. (2015). Contamination of eyewitness self-reports and the mistaken-identification problem. *Current Directions in Psychological Science*, *24*, 120–124. https://doi.org/10.1177/0963721414554394

Sparks, J. R., & Rapp, D. N. (2011). Readers' reliance on source credibility in the service of inference generation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*, 230-247. https://doi.org/10.1037/a0021331

Taber, C. S., & Lodge, M. (2006). Motivated skepticism in the evaluation of political beliefs. *American Journal of Political Science, 50,* 755-769. https://doi.org/10.1111/j.1540-5907.2006.00214.x

Tippett, C. D. (2010). Refutation text in science education: A review of two decades of research. *International Journal of Science and Mathematics Education, 8,* 951-970. https://doi.org/10.1007/s10763-010-9203-x

Unkelbach, C., & Greifeneder, R. (2018). Experiential fluency and declarative advice jointly inform judgments of truth. *Journal of Experimental Social Psychology*, *79*, 78-86. https://doi.org/10.1016/j.jesp.2018.06.010

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*, 1-8. https://doi.org/10.1162/jocn_a_00923

Weinstein, Y., McDermott, K. B., & Chan, J. C. (2010). True and false memories in the DRM paradigm on a forced choice test. *Memory, 18*, 375-384. https://doi.org/10.1080/09658211003685533

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*, 10-65. https://doi.org/10.1177/1529100616686966

Zawadzka, K., Krogulska, A., Button, R., Higham, P. A., & Hanczakowski, M. (2016). Memory, metamemory, and social cues: Between conformity and resistance. *Journal of Experimental Psychology: General, 145*, 181-199. https://doi.org/10.1037/xge0000118